

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 05-233565
 (43)Date of publication of application : 10.09.1993

(51)Int.Cl. G06F 15/16
 G06F 3/16
 G06F 13/00
 G10L 3/00

(21)Application number : 04-299468
 (22)Date of filing : 10.11.1992

(71)Applicant : FUJITSU LTD
 (72)Inventor : MATSUMOTO TATSURO

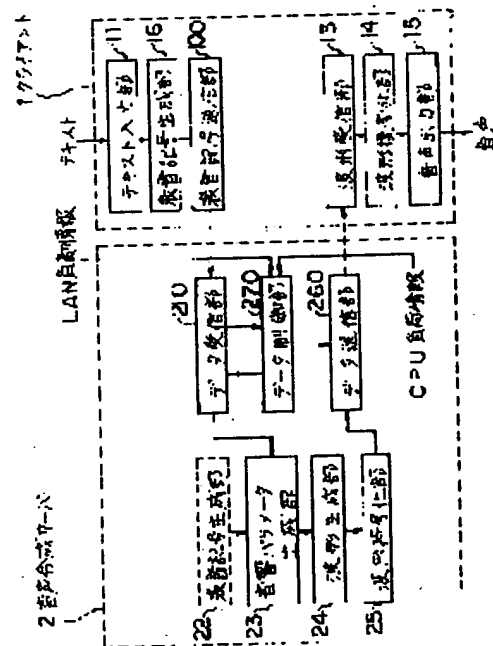
(30)Priority
 Priority number : 03295621 Priority date : 12.11.1991 Priority country : JP

(54) VOICE SYNTHESIZATION SYSTEM

(57)Abstract:

PURPOSE: To reduce the load of a voice synthesization server and also to reduce the LAN traffic.

CONSTITUTION: A data control part 270 of a voice synthesization server 2 receives the data to be voice synthesized from a client 1. If the received data is identical with the text data, the part 270 outputs the data to a phonetic symbol generating part 22 to produce the phonetic symbols. If the received data is identical with a phonetic symbol, this data is outputted to an acoustic parameter generating part 23 for generation of an acoustic parameter. If the received data is identical with an acoustic parameter, the data is outputted to a waveform generating part 24 for generation of a waveform. Then the generated voice waveforms are sent to the client 1 through a data transmission part 260. The client 1 receives the voice waveforms from the server 2 and outputs these voice waveforms as the voices through a voice output part 15.



LEGAL STATUS

[Date of request for examination] 11.12.1998
 [Date of sending the examiner's decision of rejection]
 [Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]
 [Date of final disposal for application]
 [Patent number] 3446764
 [Date of registration] 04.07.2003
 [Number of appeal against examiner's decision of rejection]
 [Date of requesting appeal against examiner's decision of rejection]
 [Date of extinction of right]

(51)Int.Cl. ⁵	識別記号	庁内整理番号	F I	技術表示箇所
G 0 6 F 15/16	3 7 0 Z	9190-5L		
3/16	3 3 0 B	8323-5B		
13/00	3 5 7 Z	7368-5B		
G 1 0 L 3/00	E	8946-5H		

審査請求 未請求 請求項の数12(全 36 頁)

(21)出願番号 特願平4-299468
 (22)出願日 平成4年(1992)11月10日
 (31)優先権主張番号 特願平3-295621
 (32)優先日 平3(1991)11月12日
 (33)優先権主張国 日本(JP)

(71)出願人 000005223
 富士通株式会社
 神奈川県川崎市中原区上小田中1015番地
 (72)発明者 松本 達郎
 神奈川県川崎市中原区上小田中1015番地
 富士通株式会社内
 (74)代理人 弁理士 大菅 義之 (外1名)

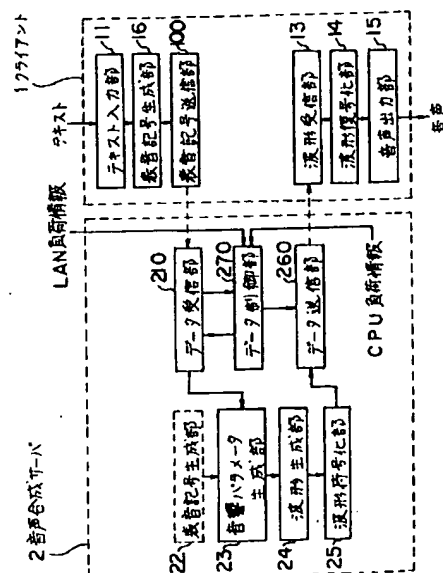
(54)【発明の名称】 音声合成システム

(57)【要約】

【目的】本発明の目的は、音声合成サーバの負荷及びLANのトラフィックを減少させることである。

【構成】クライアント1から音声合成すべきデータを受信すると、音声合成サーバ2のデータ制御部270は、受信したデータがテキストデータであれば、そのデータを表音記号生成部22へ出力して表音記号の生成を行わせ、受信したデータが表音記号であれば、そのデータを音響パラメータ生成部23へ出力して音響パラメータの生成を行わせ、受信したデータが音響パラメータであれば、そのデータを波形生成部24へ出力して音声波形の生成を行わせる。そして、データ送信部260から、生成した音声波形をクライアント1に送信する。クライアント1は、音声合成サーバ2からの音声波形を受信し、受信した音声波形を音声出力部15から音声として出力する。

本発明の実施例2を示す図



【特許請求の範囲】

【請求項1】 ローカルエリアネットワーク（LAN）上に、音声化すべきテキストデータを入力するテキスト入力部（11）と、合成音声を出力する音声出力部（15）とを有する複数のクライアント（1）と、合成音声波形を生成する波形生成部（24）を有する少なくとも1つの音声合成サーバ（2）とを備えた音声合成システムにおいて、

前記音声合成サーバ（2）は、
前記クライアント（1）から送信されるデータに基づいて音声波形を生成する波形生成部（24）と、
前記音声波形を符号化する波形符号化部（25）とを有し、

前記クライアント（1）は、
前記音声合成サーバ（2）から送信される符号化された音声波形を復号する波形復号化部（14）と、
前記波形復号化部（14）で復号された音声波形を音声として出力する音声出力部（15）とを有することを特徴とする音声合成システム。

【請求項2】 ローカルエリアネットワーク（LAN）上に、音声化すべきテキストデータを入力するテキスト入力部（11）と、合成音声を出力する音声出力部（15）とを有する複数のクライアント（1）と、音声波形を生成する波形生成部（24）を有する少なくとも1つの音声合成サーバ（2）とを備えた音声合成システムにおいて、

前記クライアント（1）は、
前記テキスト入力部11から入力されるテキストデータに基づいて表音記号を生成する表音記号生成部（16）と、

前記表音記号を前記音声合成サーバ（2）へ送信するデータ送信部（110）と、

前記音声合成サーバ（2）から送信されるデータを受信するデータ受信部（120）とを備え、

前記音声合成サーバ（2）は、
前記クライアント（1）から送信されるデータを受信するデータ受信部（210）と、

前記データ受信部（210）で受信されたテキストデータに基づいて表音記号を生成する表音記号生成部（22）と、

前記表音記号に基づいて音響パラメータを生成する音響パラメータ生成部（23）と、

前記音響パラメータに基づいて音声波形を生成する波形生成部（24）と、

前記データ受信部（210）において、テキストデータを受信した場合には、受信したテキストデータを前記表音記号生成部（22）へ出力させ、

表音記号を受信した場合には、受信した表音記号を前記音響パラメータ生成部（23）へ出力させる選択制御を行うデータ制御部（270）と、

前記波形生成部（25）で生成された音声波形を前記クライアント（1）へ送信するデータ送信部（260）とを備えることを特徴とする音声合成システム。

【請求項3】 前記クライアント（1）は、前記表音記号生成部（16）で生成される表音記号に基づいて音響パラメータを生成する音響パラメータ生成部（17）を備え、

前記音声合成サーバ（2）は、

前記データ受信部（210）において、テキストデータを受信した場合には、受信したテキストデータを前記表音記号生成部（22）へ出力させ、

表音記号を受信した場合には、受信した表音記号を前記音響パラメータ生成部（23）へ出力させ、

音響パラメータを受信した場合には、受信した音響パラメータを前記波形生成部（24）へ出力させる選択制御を行うデータ制御部（270）を備えることを特徴とする請求項2記載の音声合成システム。

【請求項4】 前記クライアント（1）は、前記データ受信部（120）で受信される音響パラメータに基づいて音声波形を生成する波形生成部（18）を備え、

前記音声合成サーバ（2）は、音声合成すべきデータを送信したクライアント（1）が波形生成部（18）を有する場合に、少なくとも自己の負荷状況及びネットワークの負荷状況に応じて、前記データ送信部（260）から音響パラメータを送信するか、音声波形を送信するかを選択制御するデータ制御部（270）を備えることを特徴とする請求項2又は3記載の音声合成システム。

【請求項5】 前記クライアント（1）は、さらにデータ制御部（130）を有し、該データ制御部（130）

は前記音声合成サーバ（2）からの指示又は前記音声合成サーバ（2）のデータ制御部（270）から送信される音声合成サーバ（2）の負荷情報及びLANの負荷情報に基づいて、前記データ送信部（110）からテキストデータを送信するか、表音記号を送信するか、音響パラメータを送信するかを選択制御することを特徴とする請求項3又は4記載の音声合成システム。

【請求項6】 前記音声合成サーバ（2）は、波形生成部（24）で生成される合成音声波形を符号化する波形符号化部（25）を有し、前記クライアント（1）は、
符号化された音声波形を復号する波形復号化部（15）を備えることを特徴とする請求項2、3、4又は5記載の音声合成システム。

【請求項7】 ローカルエリアネットワーク（LAN）上に、音声合成要求を行う複数のクライアント（1）と、前記クライアント（1）から送信されるデータに基づいて音声波形を生成し、生成した音声波形を前記クライアント（1）に送信する少なくとも1つの音声合成サーバ（2）とを備えた音声合成システムにおいて、

前記クライアント（1）がD/A変換機能を有しない場合に、D/A変換機能を有する他のクライアント（1）

を音声波形の送信先として指定し、指定したクライアント(1)から音声出力させることを特徴とする音声合成システム。

【請求項8】 ローカルエリアネットワーク(LAN)上に、音声合成要求を行う複数のクライアント(1)と、前記クライアント(1)から送信されるデータに基づいて合成音声波形を生成し、生成した音声波形をクライアント(1)に送信する少なくとも1つの音声合成サーバ(2)とを備えた音声合成システムにおいて、前記クライアント(1)がD/A変換機能を有しない場合に、特定の電話機の電話番号データを前記音声合成サーバ(2)に送信し、該電話機を音声波形の送信先として指定し、該電話機から音声出力させることを特徴とする音声合成システム。

【請求項9】 音声合成すべきテキストデータを入力するクライアント(1)と、音質又は言語毎に専用化され、前記クライアント(1)から送信されるデータに基づいて自己の管理する音質又は言語の音響パラメータを生成し、生成した音響パラメータを送信する複数の音響パラメータ生成サーバ(2a、2b)と、前記音響パラメータ生成サーバ(2a、2b)から送信される音響パラメータに基づいて音声波形を生成し、生成した音声波形を前記クライアント(1)に送信する波形生成サーバ(2c)とからなることを特徴とする音声合成システム。

【請求項10】 検索すべき単語を入力する単語入力部(41)と、前記単語入力部(41)から入力された単語を送信する単語送信部(42)と、辞書情報及び音声波形を受信する辞書情報/波形受信部(43)と、前記辞書情報を表示する辞書情報表示部(45)と、前記音声波形を音声として出力する音声出力部(15)とからなるクライアント(4)と、前記クライアント(4)から検索要求のあった単語の辞書情報を単語辞書(53)から検索する辞書検索部(52)と、検索された辞書情報の中の表音記号を送信する表音記号送信部(55)と、音声波形を受信すると共に、受信した音声波形と前記辞書検索部(52)で検索された辞書情報とを前記クライアント(4)に送信するデータ送受信部(57)とからなる辞書検索サーバ(5)と、前記辞書検索サーバ(5)から送信される表音記号を受信する表音記号受信部(220)と、前記表音記号に対応する音声波形を生成する波形生成部(24)と、前記波形生成部(24)で生成された音声波形を前記辞書検索サーバ(5)に送信する波形送信部(26)と

らなる音声合成サーバ(2)とを有することを特徴とする音声合成システム。

【請求項11】 スケジュール情報と音声波形を受信するスケジュール/波形受信部(63)と、前記スケジュール/波形受信部(63)で受信された音声波形を音声として出力する音声出力部(15)とからなるクライアント(6)と、複数のユーザのスケジュール情報を記憶するスケジュールデータベース(73)と、前記スケジュールデータベース(73)から所定の条件のスケジュール情報を検索するスケジュール検索部(72)と、検索した前記スケジュール情報を送信するスケジュール送信部(74)とからなるスケジュール管理サーバ(7)と、前記スケジュール管理サーバ(7)から送信されるスケジュール情報を受信するテキスト受信部(21)と、前記テキスト受信部(21)で受信されたスケジュール情報に対応する音声波形を生成する波形生成部(24)と、前記音声波形を前記クライアント(6)又はスケジュール管理サーバ(7)に送信する波形送信部(26)とからなる音声合成サーバ(2)とを有することを特徴とする音声合成システム。

【請求項12】 前記音声合成サーバ(2)の前記波形送信部(26)は、生成された音声波形を前記スケジュール管理サーバ(7)に送信し、前記スケジュール管理サーバ(7)のスケジュール送信部(74)は、前記音声合成サーバ(2)から送信される音声波形を受信すると共に、受信した音声波形と前記スケジュール検索部(72)で検索されたスケジュール情報とを前記クライアント(6)に送信することを特徴とする請求項1記載の音声合成システム。

【発明の詳細な説明】

【0001】

【産業上の利用分野】 本発明は、ローカルエリアネットワーク(LAN)に、複数のクライアントと少なくとも1つの音声合成サーバとが接続された音声合成システムに関する。

【0002】

【従来の技術】 LAN上にクライアントと音声合成サーバを設置し、クライアントの要求に従ってテキストの合成音声を生成し、クライアントに転送するシステムが丸川、砂川によって提案されている(丸川、砂川：第15回UNIXシンポジウム・プロシーディングズ 96～104頁)。このようなシステムの基本構成は図27に示すように、クライアント1に、テキスト入力部11、テキスト送信部12、波形受信部13、音声出力部15を設け、音声合成サーバ2に、クライアント1側のテキスト送信部12から送信されるテキストデータを受信するテ

5

キスト受信部21、表音記号生成部22、音響パラメータ生成部23、波形生成部24、波形生成部24で生成された音声合成波形をクライアント1に送信する波形送信部26とを設けた構成である。

【0003】そして、クライアント1側のテキスト入力部11からテキストデータが入力されると、テキスト送信部12が、当該テキストデータを音声合成サーバ2に送信する。音声合成サーバ2は、テキスト送信部12から送信されたテキストデータをテキスト受信部21で受信し、表音記号生成部22が、当該テキストデータがどのように発音されるのかを示す表音記号列に変換する。さらに、音響パラメータ生成部23は、表音記号列を音響パラメータ時系列に変換し、波形生成部24がその音響パラメータ時系列に基づいて音声波形を生成する。そして、波形送信部26は、生成された音声波形をクライアント1に送信する。

【0004】クライアント1側では、音声波形を波形受信部13で受信し、音声出力部15が音声波形を音声として再生する。

【0005】

【発明が解決しようとする課題】上述した従来の音声合成システムでは、クライアント1と音声合成サーバ2との間で音声データ（合成音声波形）を直接転送するため、LANのトラフィックが増大するという問題点があった。

【0006】また、従来の音声合成システムでは、サーバ/クライアントの持つ資源（機能）に関わらず、サーバ/クライアント間の通信は固定された種類のデータで行われている為に、クライアント1の資源が有効に利用されないという問題点があった。すなわち、図27のシステムでクライアント1が表音記号の生成機能を有する場合でも、クライアント1から音声合成サーバ2に送信されるデータは、常にテキストデータであり、クライアント1の機能を有効に利用していなかった。

【0007】ところで、クライアント1の中にはD/A変換機能を持たないものもあり、このようなクライアント1のユーザは、音声合成サーバ2から送られて来るデジタルデータを再生できないので、従来は、D/A変換機能を持ったクライアントしか音声データを受け取ることができないという問題点があった。

【0008】また、上述したローカルエリアネットワーク上に複数の単語情報を記憶した辞書検索サーバを設け、クライアント1のユーザから特定の単語の検索要求があったとき、辞書検索サーバが該当する単語の意味、発音記号を検索し、それらの情報をクライアントに転送して表示する辞書検索システムが知られている。単語を検索する場合、その単語の意味と同時に音声による発音が分かると便利であるが、従来そのようなシステムは存在しなかった。

【0009】また、ユーザが入力したスケジュールデー

6

タを記憶しておいて、スケジュール日時となったとき、メッセージ等でユーザに知らせるスケジュール管理システムが知られている。このようなシステムで、スケジュール内容を音声で知らせることができればさらに便利であるが、従来そのようなシステムは存在しなかった。

【0010】本発明の第1の目的は、LANに送出されるデータ量を減らし、LANのトラフィックを減少させることである。本発明の第2の目的は、音声合成サーバ及びクライアントの持つ機能に応じて両者の間を伝送されるデータの種類を変化させ、音声合成サーバの負荷及びLANのトラフィックを減少させることである。

【0011】本発明の第3の目的は、D/A変換機能を持たないクライアントのユーザが、合成音声を聞くことのできる音声合成システムを提供することである。本発明の第4の目的は、単語の発音あるいはスケジュールの内容を音声で知らせることのできるシステムを提供することである。

【0012】

【課題を解決するための手段】図1の原理図1に対応する発明は、ローカルエリアネットワーク（LAN）上に、音声化すべきテキストデータを入力するテキスト入力部11と、合成音声出力する音声出力部15とを有する複数のクライアント1と、合成音声波形を生成する波形生成部24を有する少なくとも1つの音声合成サーバ2とを備えた音声合成システムを前提にする。

【0013】図1の原理図1において、音声合成サーバ2の波形生成部24は、クライアント1から送信されるデータに基づいて音声波形を生成する。波形符号化部25は、波形生成部24で生成される音声波形を符号化する。

【0014】クライアント2の波形復号化部14は、音声合成サーバ2から送信される符号化された音声波形を復号する。音声出力部15は、復号された音声波形を音声として出力する。

【0015】図2の原理図2において、クライアント1の表音記号生成部16は、テキスト入力部11から入力されるテキストデータに基づいて表音記号（表音記号列）を生成する。ここで、表音記号とは、入力されたテキストデータの発音、アクセント、イントネーションなどを示す記号である。

【0016】音響パラメータ生成部17は、上記表音記号に基づいて音響パラメータ（音響パラメータ時系列）を生成する。データ送信部110は、クライアント1がテキスト入力部11と表音記号生成部16とを有する場合には、表音記号を音声合成サーバ2に送信し、クライアント1がテキスト入力部11と、表音記号生成部16と、音響パラメータ生成部17とを有する場合には、音響パラメータを音声合成サーバ2に送信する。

【0017】データ受信部120は、音声合成サーバ2から出力される音声波形を受信する。音声合成サーバ2

のデータ受信部210は、クライアント1から送信されるデータを受信する。

【0018】表音記号生成部22は、データ受信部210で受信されたテキストデータに基づいて表音記号を生成する。音響パラメータ生成部23は、表音記号に基づいて音響パラメータを生成する。

【0019】波形生成部24は、音響パラメータに基づいて合成音声波形を生成する。データ制御部270は、データ受信部210においてテキストデータを受信した場合には、受信したテキストデータをデータ受信部210から表音記号生成部22へ出力させ、表音記号を受信した場合には、受信した表音記号を音響パラメータ生成部23へ出力させ、音響パラメータを受信した場合には、受信した音響パラメータを波形生成部24へ出力させる選択制御を行う。

【0020】図3の原理図3において、クライアント1側の波形生成部18は、データ受信部120で受信する音響パラメータに基づいて合成音声波形を生成する。クライアント側のデータ制御部130は、データ受信部120で音響パラメータを受信した場合には、その音響パラメータを波形生成部18へ出力させ、合成音声波形を受信した場合には、その合成音声波形を音声出力部15へ出力させる選択制御を行う。

【0021】さらに、クライアント側のデータ制御部130は、音声合成サーバ2からの指示に従って、データ送信部110から音声合成サーバ2へテキストデータを送信するか、表音記号を送信するか、音響パラメータを送信するかを選択制御する。

【0022】音声合成サーバ2側のデータ制御部270は、データ受信部210においてテキストデータを受信した場合には、受信したテキストデータを表音記号生成部22へ出力させ、表音記号を受信した場合には、受信した表音記号を音響パラメータ生成部23へ出力させ、音響パラメータを受信した場合には、受信した音響パラメータを波形生成部24へ出力させる選択制御を行う。

【0023】また、データ制御部270は、音声合成要求を行ったクライアント1が波形生成部18を有する場合には、自己の負荷状況、LANの負荷状況等に応じて、データ送信部260からクライアント1へ音声波形を送信させるか、音響パラメータを送信させるかを選択制御する。

【0024】図4の原理図4において、音響パラメータ生成サーバ2a、2bは、それぞれ音質又は言語毎に専用化されており、特定の音質又は言語の音響パラメータを生成する。

【0025】クライアント1は、音響パラメータ生成サーバ2a、2bに対して特定の音質又は言語の音響パラメータの生成を指示する。波形サーバ2cは、上記音響パラメータ生成サーバ2a、2bで生成された音響パラメータに基づいて合成音声波形を生成する。

【0026】図5において、クライアント4は、単語入力部41と単語送信部42と辞書情報/波形受信部43と音声出力部15と辞書情報表示部45とを有している。単語入力部41は、検索すべき単語が入力され、単語送信部42は、その入力された単語を辞書検索サーバ5に送信する。

【0027】辞書情報/波形受信部43は、辞書情報及び単語の音声波形を受信する。辞書情報表示部45は、辞書情報/波形受信部43で受信した単語の辞書情報を表示し、音声出力部15は、受信した音声波形を音声として出力する。

【0028】辞書検索サーバ5は、単語辞書53と辞書検索部52と表音記号送信部55とデータ送受信部57とを有している。辞書検索部52は、クライアント4から検索要求のあった単語の辞書情報を単語辞書53から検索する。

【0029】表音記号送信部55は、辞書検索部52により検索された辞書情報の中の表音記号を音声合成サーバ2に送信する。データ送受信部57は、音声合成サーバ2から送信される音声波形を受信すると共に、受信した音声波形と辞書検索部52により検索された辞書情報とをクライアント4に送信する。

【0030】音声合成サーバ2は、表音記号受信部220と、音響パラメータ生成部23と、波形生成部24と、波形送信部26とを有する。表音記号受信部220は、辞書検索サーバ5から送信される表音記号を受信する。波形生成部24は、受信した表音記号に対応する音声波形を生成し、波形送信部26は、生成された音声波形を辞書検索サーバ5に送信する。

【0031】図6において、クライアント6は、スケジュール/波形受信部63と音声出力部15とスケジュール表示部62とを有している。スケジュール/波形受信部63は、スケジュール情報及び音声波形を受信する。スケジュール表示部62は、スケジュール/波形受信部63で受信したスケジュール情報を表示し、音声出力部15は、受信した音声波形を音声として出力する。

【0032】スケジュール管理サーバ7は、スケジュールデータベース73とスケジュール検索部72とスケジュール送信部74とを有する。スケジュールデータベース73は、複数のユーザのスケジュール情報を記憶する。

【0033】スケジュール検索部72は、所定の条件のスケジュール情報をスケジュールデータベースから検索する。スケジュール送信部74は、検索されたスケジュール情報を音声合成サーバ2に送信する。

【0034】音声合成サーバ2は、テキスト受信部21と、表音記号生成部22と、音響パラメータ生成部23と、波形生成部24と、波形送信部26とを有している。テキスト受信部21は、スケジュール管理サーバ7から送信されるスケジュール情報を受信する。波形生成

部24は、受信されたスケジュール情報に対応する音声波形を生成し、波形送信部26は、その音声波形をスケジュール管理サーバ7又はクライアント6に送信する。

【0035】

【作用】図1の原理図1に対応する発明では、音声合成サーバ2の波形生成部24がクライアント1から送信されたデータに対応する音声波形を生成し、その音声波形を符号化してクライアント1に送信する。

【0036】クライアント1側では、受信した音声波形を復号化部14で元の音声波形に復号し、音声出力部15から音声として出力する。この場合、音声合成サーバ2からLANには圧縮されたデータ（符号化された音声波形）が送出されるので、LANを伝送されるデータ量が少なくなり、LANのトラヒックが減少する。

【0037】次に、図2の原理図2に対応する発明では、クライアント1は、音声出力部15とデータ送信部110とデータ受信部120の他に、(i)テキスト入力部11のみを有する場合と、(ii)テキスト入力部11と表音記号生成部16とを有する場合と、(iii)テキスト入力部11と表音記号生成部16と音響パラメータ生成部17とを有する場合がある。

【0038】クライアント1が(i)の構成を有する場合には、データ送信部110から音声合成サーバ2にテキストデータを送信し、(ii)の構成を有する場合には、データ送信部110から音声合成サーバ2に表音記号を送信し、(iii)の構成を有する場合には、データ送信部110から音声合成サーバ2に音響パラメータを送信する。

【0039】音声合成サーバ2側では、クライアント1からテキストデータを受信した場合には、受信したテキストデータを表音記号生成部22へ出力して表音記号の生成を行い、表音記号を受信した場合には、受信した表音記号を音響パラメータ生成部23へ出力して音響パラメータの生成を行い、音響パラメータを受信した場合には、受信した音響パラメータを波形生成部24へ出力して合成音声波形の生成を行う。

【0040】この場合、クライアント1は、自己の有する機能に応じたデータを音声合成サーバ2に送信するので、クライアント1が表音記号生成機能、あるいは音響パラメータ生成機能を有する場合には、音声合成サーバ2の表音記号の生成処理、あるいは音響パラメータの生成処理等が不要となり、音声合成サーバ2の負荷を軽減することができる。

【0041】次に、図3の原理図3に対応する発明では、クライアント1は、原理図2の発明の構成に加えデータ制御部130を有する。この場合、クライアント1は、(i)テキスト入力部11のみを有する場合と、(ii)テキスト入力部11と表音記号生成部16とを有する場合と、(iii)テキスト入力部11と表音記号生成部16と音響パラメータ生成部17とを有する場合とがある。

【0042】音声合成サーバ2のデータ制御部270は、クライアント1から音声データの送信要求があったとき、音声合成サーバ2の負荷状況、LANの負荷状況及び要求のあったクライアント1の有する機能に基づいて、クライアント1から音声合成サーバ2に送信すべきデータの種類の決定し、決定したデータの種類のクライアント1のデータ制御部130に指示する。あるいは、音声合成サーバ2のデータ制御部270からクライアント1のデータ制御部130に、音声合成サーバ2の負荷状況、LANの負荷状況に関する情報を通知し、クライアント1のデータ制御部130自身が送信すべきデータの種類の決めることもできる。

【0043】例えば、音声合成サーバ2の負荷が大きい場合には、音声合成サーバ2のデータ制御部270は、クライアント1のデータ制御部130にクライアント1の持つ機能に応じた種類のデータを送信するように指示する。そして、クライアント1のデータ制御部130は、音声合成サーバ2から指示された種類のデータの送信をデータ送信部110に指示する。

【0044】具体的には、クライアント1が上記(ii)の構成を有する場合には、クライアント1のデータ制御部130は、表音記号の送信をデータ送信部110に指示し、クライアント1が上記(iii)の構成を有する場合には、音響パラメータの送信を指示する。

【0045】一方、音声合成サーバ2の負荷が小さい場合には、クライアント1側の処理を音声合成サーバ2側で行うことができるので、音声合成サーバ2のデータ制御部270は、それぞれのクライアント1が持っている機能以下のデータを送信するように指示する。

【0046】例えば、音声合成サーバ2の負荷が小さく、処理能力に余裕がある場合には、音声合成サーバ2のデータ制御部270は、クライアント1が表音記号の生成機能を有する場合（上記(ii)の場合）でもクライアント1のデータ制御部130にテキストデータの送信を指示する。

【0047】また、音声合成サーバ2のデータ制御部270は、クライアント1が音響パラメータまで送信できる機能を有する場合（上記(iii)の場合）でも、クライアント1のデータ制御部130に表音記号（又はテキストデータ）の送信を指示する。

【0048】このように、音声合成サーバ2の負荷が小さい場合に、クライアント1で行う表音記号の生成処理、あるいは音響パラメータの生成処理を音声合成サーバ2側で行うことで、クライアント1の負荷を減らすことができる。この場合、クライアント1から音声合成サーバ2へ送信されるデータは、表音記号の代わりに、よりデータ量の少ないテキストデータ、あるいは音響パラメータの代わりに、よりデータ量の少ない表音記号（又はテキストデータ）となるので、LANのトラヒックを減少させることもできる。

【0049】さらに、原理図3に対応する発明のクライアント1は、波形生成部18を有する場合がある。この場合、音声合成サーバ2は、自己の負荷状況、LANの負荷状況及びクライアント1の負荷状況に応じて合成音声波形、音響パラメータの何れかを選択してクライアント1に送信する。

【0050】例えば、音声合成サーバ2の負荷が小さい場合には、音声合成サーバ2は、音声波形の生成まで行い、データ送信部260からその音声波形をクライアント1に送信する。

【0051】一方、音声合成サーバ2の負荷が大きい場合には、波形生成部24での音声波形の生成は行わず、音響パラメータをクライアント1に送信する。また、LANのトラフィックが大きいときには、音声波形よりデータ量の少ない音響パラメータを送信する。

【0052】このように、音声合成サーバ2の負荷の状況、LANの負荷状況、クライアントの負荷状況に応じて音声波形、音響パラメータの何れかを選択してクライアント1に送信することにより、音声合成サーバ2の負荷を軽くすることができる。また、音声波形の替わりによりデータ量の少ない音響パラメータを送信することで、LANの負荷を減少させることができる。

【0053】図4の原理図に対応する発明では、クライアント1が特定の音質又は言語の音響パラメータの生成を指示すると、該当する音響パラメータサーバ2a（又は2b）が、クライアント1から送信されるテキストデータに基づいて音響パラメータを生成し、その音響パラメータを波形生成サーバ2cに送信する。波形生成サーバ2cは、受信した音響パラメータに基づいて音声波形を生成し、その音声波形をクライアント1に送信する。これにより、クライアント1からは、指定した音質又は言語でテキストデータが音声出力される。

【0054】図5の原理図に対応する発明では、辞書検索部52は、クライアント4から検索要求のあった単語の辞書情報を単語辞書53から検索し、検索した辞書情報の中の表音記号を音声合成サーバ2に送信する。音声合成サーバ2は、その表音記号に対応する音声波形を生成し、生成した音声波形を辞書検索サーバ5に送信する。

【0055】辞書検索サーバ5は、辞書検索部52が検索した単語の辞書情報と、音声合成サーバ2から受信した音声波形とをクライアント4に送信する。クライアント4は、単語の辞書情報を辞書情報表示部45に表示すると共に、単語の発音を音声出力部15から音声で出力する。

【0056】これにより、ユーザがクライアント4に特定の単語の検索要求を行ったとき、その単語の意味が表示されると共に、その発音が音声で出力されるので、ユーザは単語の意味とその発音を正確に知ることができる。

【0057】次に、図6の原理図6に対応する発明では、スケジュール管理サーバ7は、スケジュールデータベース73に記憶されているスケジュール情報を検索し、例えばスケジュール時刻が現在時刻と一致するスケジュール情報を取り出す。そして、そのスケジュール情報を音声合成サーバ2に送信する。音声合成サーバ2は、テキストデータの形で受信するスケジュール情報に対応する音声波形を生成し、その音声波形をスケジュール管理サーバ7に送信する。

10 【0058】スケジュール管理サーバ7は、音声波形をスケジュール情報と共にクライアント6に送信する。クライアント6は、スケジュール情報をスケジュール情報表示部62に表示し、スケジュール内容を示す音声を音声出力部15から出力する。なお、音声合成サーバ2で生成した音声波形を直接クライアント6に送信するようにしても良い。

【0059】これにより、スケジュール情報がディスプレイ等に表示されると共に、音声で出力されるので、ユーザに対するスケジュールの通知をより確実に行うことができる。

20 【0060】

【実施例】以下、本発明の実施例を図面を参照しながら説明する。図7は、本発明の実施例1の音声合成システムの構成図である。この実施例は、少なくとも1つの音声合成サーバ2と複数のクライアント1とが接続されてLANが形成されている場合を示している。

【0061】クライアント1には、音声合成すべきメッセージをテキストデータとして入力するテキスト入力部11、このテキスト入力部11から入力されたテキストデータを音声合成サーバ2に送信するテキスト送信部12、音声合成サーバ2から送信される符号化された音声波形を受信する波形受信部13、受信した音声波形を復号する波形復号化部14、復号された音声ディジタルデータをアナログデータに変換するDA変換部151、アナログデータから高周波成分を除去するローパスフィルタ152が設けられ、ローパスフィルタ152にスピーカ153が接続されている。

【0062】一方、音声合成サーバ2には、クライアント1のテキスト送信部12から送信されてくるテキストデータを受信するテキスト受信部21、受信したテキストデータ内の略字や数字の読みの決定を行う前処理部220、単語辞書224からの情報を参照し形態素解析および構文解析を行う言語処理部221、テキストデータがどのように発音されるかを示す発音記号列に変換する発音記号生成部222が設けられている。なお、表音記号＝発音記号＋韻律記号である。

【0063】また、韻律規則225を参照し、かつ言語処理部221及び発音記号生成部222の結果に基づきイントネーションやアクセントを示す韻律記号を生成する韻律記号生成部223、音響パラメータを格納したパ

ラメータファイル231を参照しつつ発音記号列及び韻律記号列に基づいて音響パラメータ時系列を生成する音響パラメータ生成部23が設けられている。

【0064】さらに、音響パラメータ時系列に基づいて音声波形を生成する波形生成部24、生成された音声波形を符号化する波形符号化部25、符号化された音声波形を送信する波形送信部26が設けられている。

【0065】以下、各部を必要に応じて詳細に説明する。

《前処理部》前処理部220は、当該テキストデータ内の略字や数字の読みの決定を行う部分であり、テキストデータ内に略字や記号、あるいは数字などがあった場合、それらを抽出し、抽出した略字等の読みを、図8に示す対応テーブルを参照して決定する。

《言語処理部》言語処理部221は、単語辞書224からの情報を参照し、入力テキストの形態素解析及び構文解析を行う。ここで、形態素解析とは、入力テキストを構成する意味のある最小の単位（形態素）に区切り、その品詞、読みやアクセントを決定する処理である。また、構文解析とは、入力テキストに自然なイントネーションを与えるために、形態素解析結果に基づいて文法的に意味のある単位を解析する処理である。

【0066】発音記号生成部222は、テキストがどのように発音されるかを示す発音記号列に変換する。発音記号列とは、発音を表す記号列であり、日本語においては平仮名表記の一字に相当する。もしくは、その文字の置かれる環境に応じて細分化したものである。

《韻律記号生成部》韻律記号生成部223は、韻律規則225を参照して言語処理部221の解析結果及び発音記号生成部222で生成される発音記号に基づきイントネーションやアクセントを示す韻律記号を生成する。

【0067】ここで参照する韻律規則225とは、言語処理部221において得られた構文情報に基づいて、意味の切れ目に休止記号やイントネーションの再立ち上がり制御する記号を挿入するときの規則を示す。また、単語と付属語や単語同士が接続することによって、単語本来を持つアクセントの位置が変化するときの規則を示す。

【0068】次に、音響パラメータ生成部23は、パラメータファイル231を参照し、発音記号列及び韻律記号列に基づいて音響パラメータ時系列を生成する。音響パラメータとは、実際の音声データを何らかの方法によって合成単位（例えば音素や音節）毎に情報圧縮したデータである。音響パラメータとしては、情報圧縮の方式の違い（音声生成過程のモデル化の違い）によって、PARCOR（LPC）、LSP、フォルマント等がある。

【0069】例えば、フォルマント（声道の共振周波数）の場合は、発音記号列に基づきフォルマント周波数、フォルマント帯域幅を指定することにより声道伝達

関数をモデル化したデジタル・フィルタを構成し、そのフィルタを、韻律記号列に基づいて生成される音源信号により駆動し音声波形を生成する。

【0070】波形生成部24は、音響パラメータ時系列に基づいて音声波形を生成し、波形符号化部25は生成された音声波形を情報圧縮する。情報圧縮の手法としては、ADM、ADPCM、APC、SBC、ATC、PARCOR、MPC、VQなどの符号化方式を用いる。

【0071】ここで、上記の音響パラメータ生成部23と波形生成部24とで採用される情報圧縮の手法について説明する。ADMとは、Adaptive Delta Modulationの略で、ADMは予測符号化の最も簡単なもので、入力信号と予測値との差を1ビット（正、負各1レベル）で符号化する方法である。ADMでは、1標本値当たり2量子化レベルしか許されないため、入力信号のサンプリング周波数を高くすることにより、隣接サンプル間の差分を小さくして符号化している。この差分の振幅があまり変化しない区間ではステップサイズを小さくし、大きく変化するときはステップサイズを大きくして適応量子化を行う。ADM符号化の構成を図9に示す。

【0072】ADPCMは、Adaptive Differential Pulse Code Modulationの略で、過去の入力信号から現在の入力信号を予測し、それと現在の入力信号との差分だけを量子化した後符号化する方法であり、入力信号レベルが急激に変化したときは量子化雑音が大きくなるため、量子化のステップ幅をこれに対応して増減する適応量子化を行う。予測方式には、予測器の予測係数が固定された固定予測方式と、信号に適応して過去の入力信号から予測誤差を最小とするように予測係数を決める適応予測方式とがある。後者は前者に比べて複雑なアルゴリズムを要するが、高品質な符号化特性が得られる。ADPCM符号化の構成を図10に示す。

【0073】APCとは、Adaptive Predictive Codingの略で、入力信号をブロック単位に取り込んで線形予測分析することにより予測係数を求め、これに基づき予測信号を発生する方法であり、予測信号と入力信号との差分を量子化して符号化するとともに、予測係数と量子化ステップ幅を補助情報として伝送する。APC符号化は逐次的に予測するADM、ADPCMとは異なり、ブロック全体のサンプルに対して最適な予測係数を求めて線形予測を行うため、予測利得が大きくなり、符号化品質が向上する。その構成を図11に示す。

【0074】SBCとは、Sub-Band Codingの略で、音声信号を複数の周波数帯域に分割し、各帯域信号をナイキスト速度でダウンサンプリングするとともにベースバンド信号に変換し、ADPCMなどにより符号化する方法である。その構成を図12に示す。各帯域の符号化において、音声のエネルギーが集中している帯域に、より多くの量子化ビット数を割り当てることにより、総合的な音声品質の向上を可能とする。また、量子化雑音が分

割された帯域にのみ発生し、他の帯域に影響を与えないようにできるという利点がある。SBCでは、2分割を単位にして帯域分割を行うことを基本とし、折り返し成分がキャンセルできるQMF (Quadrature Mirror Filter) を用いるのが一般的である。

【0075】ATCとは、Adaptive Transform Codingの略で、音声がほぼ正常と考えられる20ms程度を1ブロック(フレーム)とし、そのブロックの音声波形信号を周波数領域に直交変換した量を符号化する方法である。復号では、周波数領域の量を逆直交変換することにより1ブロック分の音声波形信号を再現する。直交変換としてDCT (Discrete Cosine Transform) を用い、適応的にビット割当てを行う方式をATCと呼ぶことが多い。ATC符号化は、人間の聴覚の周波数分析機能との対応づけや、量子化雑音の周波数領域の制御が容易であるという特徴を有する。基本的な構成を図13に示す。

【0076】PARCORとは、PARTIAL auto CORrelationの略で、人間の音声は、声帯の振動、あるいは口を狭めることで発生する乱流による音源信号が、音道、口腔、鼻腔などにおける共振により固有のスペクトル分布を有する音響信号に変換されて生成されることに着目し、音声信号から駆動音源情報およびスペクトル分布情報を抽出(分析)し、これらのパラメータを用いて音声信号を復元(合成)する音声分析合成方法である。図14に示すように、有音声の場合は声帯振動を模擬するパルス発生器、無声音の場合は乱流を模擬する白色雑音発生器により、予測フィルタを駆動することによって音声を合成する。

【0077】本方式では、スペクトル情報として予測係数の代わりにPARCOR係数を用いている。このPARCOR係数は、伝送効率が優れ、パラメータ符号化時の量子化誤差に対して予測フィルタの安定性確保が可能である。音源情報は、音声信号を逆フィルタ(予測フィルタと逆の特性を有するフィルタ)に通すことによって得られる予測残差信号から抽出する。具体的には、残差信号の自己相関係数が最大となるサンプル間隔からパルス周期(ピッチ)を求め、このときの係数値と1次のPARCOR係数値を用いて有/無声判定を行う。符号化情報として、一定のフレーム周期(20ms程度)毎にPARCOR係数、ピッチ周期、有/無声判定および残差電力情報を量子化して伝送する。

【0078】また、その後開発されたLSP方式は、PARCOR係数の代わりにLSP (Line Spectrum Pair) 係数を用いるもので、音声のスペクトル情報抽出方法の簡単化、およびパラメータ符号化時の情報圧縮率の点で著しく改善されている。

【0079】MPCとは、Multi-Pulse Codingの略で、線形予測(LPC)分析合成系(ボコーダ)において、パルスと雑音による音源のモデル化を避け、音源を有声音、無声音にかかわらず複数のパルスによって表現し、

これによってLPC合成フィルタを駆動する方式である。MPC符号化のポイントは、いかにして複数のパルスの振幅と時間位置を最適に設定するかであり、図15に示すようなA-b-S (Analysis-by-Synthesis)の原理による方法が用いられている。まず、予め音声信号を20ms程度の長さのブロック(フレーム)に切り出し、その区間についてのスペクトル包絡を線形予測分析によって推定する。次に、このフレーム内の音声信号について音源のマルチパルスを決定する。具体的には、すでにある数のパルスが決定されているとすると、このマルチパルスは、推定されたスペクトル包絡に対応するLPC合成フィルタを通して合成音声に変換される。この合成音声と実際の原音声信号の誤差を計算し、聴覚特性による重み付けを行ってから、平均二乗誤差が最小になるように、新しいパルスの振幅と時間位置を決定する。

【0080】VQとは、Vector Quantizationの略で、波形符号化や分析合成系において、波形やスペクトル包絡パラメータを各サンプル値ごとに量子化せず、複数の値の組(ベクトル)をまとめて1つの符号で表現する量子化方法である。ベクトル量子化は初めLPCパラメータの高効率な量子化方法として提案され、その後、波形の量子化にも適用されるようになった。図16のベクトル量子化による符号化では、標準化した波形を一定の時間について切り出し、その区間の波形パターンを1つの符号で表現する。このため、予めクラスタリングの手法を用いて種々の波形パターンを蓄えておき、それぞれに符号を与えておく。この符号とパターンとの対応を示す表を符号帳(codebook)と呼ぶ。入力波形に対して、一定の時間区間ごとに符号帳の各パターンと比較(パターンマッチング)し、最も類似度の高い波形の符号でその区間の波形を表現する。

【0081】次に、図7に戻り、実施例1の音声合成処理を説明する。ユーザによりクライアント1のテキスト入力部11からテキストデータが入力されると、入力されたテキストデータがテキスト送信部12から音声合成サーバ2へ送信される。

【0082】音声合成サーバ2では、テキスト受信部21が、そのテキストデータを受信し、前処理部220が、当該テキストデータ内の略字や数字の読みの決定を行う。さらに、言語処理部221は、単語辞書224の情報を参照し、テキストデータの形態素解析及び構文解析を行う。発音記号生成部222は、テキストデータがどのように発音されるかを示す発音記号列に変換し、韻律記号生成部223は、言語処理部221及び発音記号生成部222の結果に基づきイントネーションやアクセントを示す韻律記号列を生成する。

【0083】次に、音響パラメータ生成部23は、パラメータファイル231を参照し、発音記号列及び韻律記号列に基づいて音響パラメータ時系列を生成する。波形生成部24は、音響パラメータ時系列に基づいて音声波

形を生成する。波形符号化部25は生成された音声波形を情報圧縮する。波形送信部26は情報圧縮された合成波形をクライアント1の波形受信部13へ送信する。

【0084】クライアント1側では、波形受信部13がサーバ2からの情報圧縮された波形を受信し、波形復号化部14が符号化された音声波形を復号する。DA変換部151はデジタル音声波形をアナログ音声波形に変換する。そして、そのアナログ音声信号がローパスフィルタ152を通してスピーカ152から出力される。

《実施例2及び実施例3》図17は、クライアント1が表音記号生成機能を有する本発明の実施例2のシステム構成図であり、図18は、クライアント1が音響パラメータ生成機能を有する実施例3のシステム構成図である。

【0085】図17に示す実施例2のクライアント1には、テキスト入力部11と、表音記号生成部16と、データ送信部110としての表音記号送信部100と、データ受信部120としての波形受信部13と、波形復号化部14及び音声出力部15とが設けられている。

【0086】また、図18に示す実施例3のクライアント1には、テキスト入力部11と、表音記号生成部16と、音響パラメータ生成部17と、データ送信部110としての音響パラメータ送信部101と、データ受信部120としての波形受信部13、波形復号化部14及び音声出力部15とが設けられている。

【0087】なお、音声合成サーバ2は、いずれの実施例においても、データ受信部210と、表音記号生成部22と、音響パラメータ生成部23と、波形生成部24と、波形符号化部25と、データ送信部260と、データ制御部270とを備えている。

【0088】音声合成サーバ2のデータ制御部270は、データ受信部210で受信したデータの種別を判別すると共に、LAN負荷情報（トラフィック情報）と自己のCPU負荷情報とに基づいてデータ受信部210及びデータ送信部260を制御する。また、データ制御部270は、図示しないクライアント1のデータ制御部に送信すべきデータの種別を指示する。

【0089】すなわち、データ制御部270は、データ受信部210で受信したデータの種別を判別し、その判別結果に応じて受信したデータを表音記号生成部22、音響パラメータ生成部23、波形生成部24のいずれかに選択的に出力する。

【0090】以下、実施例2及び実施例3の動作を説明する。

《実施例2の場合》図17において、ユーザによりクライアント1のテキスト入力部11からテキストデータが入力されると、表音記号生成部16は、そのテキストデータがどのように発音されるかを示す表音記号列に変換する。そして、表音記号送信部100は、その表音記号列を音声合成サーバ2に送信する。

【0091】音声合成サーバ2のデータ受信部210は、クライアント1からのデータ（表音記号列）を受信すると、データ制御部270が、データ受信部210で受信されたデータの種別を判別し、そのデータを該当する回路に出力させる。この場合、クライアント1からは送信データとして表音記号列が送られてくるので、データ制御部270は、その表音記号列を音響パラメータ生成部23に出力させる。

【0092】音響パラメータ生成部23は、データ受信部210で受信された表音記号列を音響パラメータ時系列に変換し、波形生成部24はその音響パラメータ時系列に基づいて音声波形を生成する。さらに、波形符号化部25は、生成された音声波形を符号化して情報圧縮し、データ送信部26は、その符号化された音声波形をクライアント1へ送信する。

【0093】クライアント1側の波形受信部13は、音声合成サーバ2から符号化された音声波形を受信し、波形復号化部14がその符号化された音声波形を復号する。そして、音声出力部15から音声として出力する。

《実施例3の場合》図18において、ユーザによりクライアント1のテキスト入力部11からテキストデータが入力されると、表音記号生成部16が、当該テキストがどのように発音されるかを示す表音記号列に変換する。さらに、音響パラメータ生成部17が、その表音記号列を音響パラメータ時系列に変換し、音響パラメータ送信部101が音響パラメータを音声合成サーバ2に送信する。

【0094】音声合成サーバ2側では、データ受信部210が、クライアント1からのデータ（音響パラメータ時系列）を受信し、データ制御部270が、受信したデータ（音響パラメータ）の種別を判別し該当する回路に出力させる。この場合、クライアント1からは送信データとして音響パラメータ時系列が送られてくるので、その音響パラメータ時系列を波形生成部24に送る。

【0095】波形生成部24は、音響パラメータ時系列に基づいて音声波形を生成し、波形符号化部25が、生成された音声波形を符号化して情報圧縮する。そして、データ送信部260が、符号化された音声波形をクライアント1へ送信する。

【0096】クライアント1側では、波形受信部13が音声合成サーバ2からの符号化された音声波形を受信し、波形復号化部14がその符号化された音声波形を復号する。さらに、音声出力部15が復号された音声波形を音声として出力する。

【0097】上述した実施例2及び3では、音声合成サーバ2は、クライアント1から送信されたデータが表音記号であれば、音響パラメータの生成以降の処理を行い、送信されたデータが音響パラメータであれば、音声波形の生成以降の処理を行えばよいので、音声合成サーバ2の負荷が軽減される。

【0098】なお、図17、図19には示していないが、クライアント1側にも、音声合成サーバ2と同様にデータ制御部130（図3参照）が設けられており、音声合成サーバ2のデータ制御部270が、LANの負荷状況、音声合成サーバ2の負荷状況に応じて送信すべきデータの種類のクライアント1のデータ制御部130に指示するようになっている。

【0099】例えば、音声合成サーバ2の負荷が大きい場合には、クライアント1側でより多くの処理を行う方が望ましいので、音声合成サーバ2のデータ制御部270が、クライアント1のデータ制御部130に、クライアント1が持つ機能に応じたデータの送信を指示する。

【0100】一方、音声合成サーバ2の負荷が小さいときには、クライアント1で行う処理を音声合成サーバ2で引き受けることができるので、例えば、クライアント1が表音記号の生成機能を有する場合でも、クライアント1のデータ制御部130にテキストデータの送信を指示する。また、クライアント1が音響パラメータの生成機能を有する場合でも、そのクライアント1のデータ制御部130にテキストデータ又は表音記号の送信を指示する。これにより、クライアント1の負荷が軽減される。また、この場合、クライアント1からLAN送出されるデータは、よりデータ量の少ないデータとなるので、LANのトラフィックも減少する。

【0101】なお、クライアント1と音声合成サーバ2との機能分担は、上述したように音声合成サーバ2のデータ制御部270がクライアント1のデータ制御部130に指示する方法に限らず、例えば、クライアント1のデータ制御部130が、音声合成サーバ2のデータ制御部270と通信して、CPU負荷情報やLAN負荷情報を受け取り、それらの情報に基づいてクライアント1のデータ制御部130自身が送信するデータの種類の決定するようにしても良い。

《実施例4》図19は、クライアント1が波形生成部18を有する本発明の実施例4のシステム構成図である。

【0102】図19に示すクライアント1には、テキスト入力部11と、表音記号生成部16と、データ送信部110としての表音記号送信部100と、データ受信部120としての音響パラメータ受信部102と、波形生成部18及び音声出力部15とが設けられている。

【0103】図19において、ユーザによりクライアント1のテキスト入力部11からテキストデータが入力されると、表音記号生成部16は、当該テキストがどのように発音されるかを表す表音記号列に変換し、表音記号送信部100がその表音記号列を音声合成サーバ2に送信する。

【0104】音声合成サーバ2では、データ受信部210が、クライアント1から送信されるデータ（表音記号列）を受信し、データ制御部270が、受信したデータの種類の判別する。この場合、受信したデータは表音記

号列であるので、データ受信部210は、データ制御部270の指示に従って、受信した表音記号列を音響パラメータ生成部23に送る。さらに、音響パラメータ生成部23は、その表音記号列を音響パラメータ時系列に変換する。

【0105】音声合成サーバ2からクライアント1に音声データを送信する場合、データ制御部270は、LANの負荷状況、自己のCPUの負荷状況、送信先のクライアント1の持つ機能に応じて送信すべきデータの種類の決定する。この場合、送信先のクライアント1が波形生成部18を有するので、データ制御部270は、音声合成サーバ2の負荷が大きいときには、音声波形の代わりに音響パラメータ生成部23で生成される音響パラメータ時系列をクライアント1へ送信させ、音声合成サーバ2の負荷が小さいときには、波形生成部24で音声波形を生成させ、その音声波形をクライアント1に送信させる。

【0106】クライアント1側の音響パラメータ受信部102は、音声合成サーバ2から送信される音響パラメータ時系列を受信し、波形生成部18がその音響パラメータ時系列に基づいて音声波形を生成する。最後に音声出力部15が音声波形を出力する。なお、音声合成サーバ2から音声波形が送信された場合には、音響パラメータ受信部102は、受信した音声波形をそのまま音声出力部15に出力する。

【0107】この実施例では、クライアント1が波形生成機能を有する場合、音声合成サーバ2は、音声波形を生成せずに音響パラメータ時系列をクライアント1に送信することができるので、音声合成サーバ2の負荷が軽減される。さらに、この場合、音声合成サーバ2からLANに送出されるデータは、音声波形よりデータ量の少ない音響パラメータ時系列となるのでLANのトラフィックも減少する。

【0108】次に、図20は、上述した実施例1～4のクライアント1と音声合成サーバ2とからなる音声合成システムのシステム構成図である。以下、音声合成サーバ2の負荷が大きい場合、負荷が小さい場合の動作を説明する。

【0109】音声合成サーバ2の負荷が大きい場合には、音声合成サーバ2での処理を可能な範囲でクライアント1側に負担させることが望ましいので、音声合成サーバ2のデータ制御部270は、各クライアント1の図示しないデータ制御部にそれぞれのクライアント1が持つ機能に応じたデータの送信を指示する。

【0110】従って、クライアント1から音声合成サーバ2にデータを送信する場合、テキストデータの生成機能を有するクライアント1は、テキストデータを音声合成サーバ2に送信し、表音記号の生成機能を有するクライアント1は、表音記号列を音声合成サーバ2に送信し、音響パラメータ生成機能を有するクライアント1は

音響パラメータ時系列を音声合成サーバ2に送信する。また、音声合成サーバ2からクライアント1にデータを送信する場合にも、クライアント1が波形生成機能を有するときには、音声合成サーバ2は音声波形の生成処理は行わず、音響パラメータ時系列を該当するクライアント1に送信し、クライアント1側で音声波形の生成を行わせる。

【0111】このように、クライアント1側で表音記号の生成処理、音響パラメータの生成処理、あるいは音声波形の生成処理を行わせることで、音声合成サーバ2の負荷を軽減させることができる。

【0112】一方、音声合成サーバ2の負荷が小さい場合には、音声合成サーバ2側でより多くの処理を行えるので、クライアント1が表音記号、あるいは音響パラメータの生成機能を有しているときにも、クライアント1にテキストデータ、あるいは表音記号の送信を指示する。また、クライアント1が波形生成機能を有するときでも、音声合成サーバ2が音声波形まで生成してクライアント1に送信する。これにより、クライアント1の処理負担が軽減される。また、この場合、LANに送出されるデータは、よりデータ量の少ないデータとなるので、LANのトラフィックも減少する。

【0113】上記実施例によれば、クライアント1の機能に応じて音声合成サーバ2側の処理を可変にすること、あるいはLANのトラフィック状況、音声合成サーバ2のCPU負荷状況に応じて、サーバ/クライアント間の処理分担を変えることができるので、システム全体の処理効率を高めることができる。

《実施例5》図21は、音質又は言語毎に専用化した複数の音響パラメータサーバ2a、2bと、共通化した波形生成サーバ2cとを設けた本発明の実施例5の概念図であり、図22は、図21のクライアント1、音響パラメータサーバ2a、2b及び波形生成サーバ2cの具体的な構成の一例を示す図である。

【0114】図21の音声合成システムには、男性、女性、子供などの音質あるいは言語毎に専用化した音響パラメータサーバ2a、2bが設けられている。クライアント1は、特定の音響パラメータ生成サーバ2a又は2bを指定し、あるいはテキストデータにより定まる音響パラメータ生成サーバ2a又は2bに対して音声合成要求を行い、音響パラメータ時系列を生成させる。

【0115】音響パラメータ生成サーバ2a又は2bは、生成した音響パラメータ時系列を波形生成サーバ2cに転送する。波形生成サーバ2cは、音響パラメータ生成サーバ2a又は2bから送信された音響パラメータ時系列に基づいて音声波形を生成し、その音声波形を音声合成要求を出したクライアント1に転送する。クライアント1はその音声波形を受信し音声として再生する。

【0116】具体的には、図22のクライアント1のテキスト入力部11から入力されたテキストデータが、テ

キスト送信部12から音響パラメータ生成サーバ2a（又は2b）のテキスト受信部21に送信される。

【0117】該当する音響パラメータ生成サーバ2a（又は2b）の表音記号生成部22は、受信したテキストデータに基づいて表音記号列を生成し、音響パラメータ生成部23がその表音記号列に基づいて音響パラメータ時系列を生成する。音響パラメータ生成部23で生成された音響パラメータ時系列は、音響パラメータ送信部310から波形生成サーバ2cに送信される。

10 【0118】そして、波形サーバ2cにおいて音響パラメータ時系列に基づいて音声波形が生成され、その音声波形がクライアント1に送信される。この実施例では、音質又は言語毎に専用化した音響パラメータ生成サーバ2a、2bを複数設けることで、テキストデータを異なった音質、言語で音声出力させることができる。

《実施例6》図23は、本発明の実施例6の概念図である。実施例6は、D/A変換機能を持たないクライアント1が、テキストデータに基づいて生成される合成音声

20 【0119】D/A変換機能を持たないクライアント1は、LANに接続されている音声合成サーバ2に対して音声合成したいテキストデータと音声出力したい電話機の電話番号データを送信する。

【0120】音声合成サーバ2は、テキストデータと電話番号データを受信すると、そのテキストデータに基づいて音声波形を生成し、指定された電話番号のダイヤル

30 【0121】この実施例6によれば、クライアント1がD/A変換機能を持たない場合でも、クライアント1の近くにある電話機に合成音声を出力させることで、テキストの内容を音声で聞くことができる。

《実施例7》図24は、本発明の実施例7の概念図である。この実施例は、D/A変換機能を持たないクライアント1が、テキストデータに基づいて生成される合成音声を、D/A変換機能を有する別のクライアントから出力させる場合の例である。

40 【0122】D/A変換機能を持たないクライアント1aは、LANに接続されている音声合成サーバ2に音声合成したいテキストデータと合成音声を出力させたいクライアント1bのアドレスデータとを送出する。音声合成サーバ2は、クライアント1aからテキストデータと特定のクライアント1bを指定するアドレスデータとを受信したなら、テキストデータに対応する音声波形を生成する。さらに、音声合成サーバ2は、指定されたアドレスのクライアント1bに、生成した音声波形を送信する。クライアント1bは、受信した音声波形から音声を再生する。

【0123】この実施例7によれば、D/A変換機能を持たないクライアント1aのユーザが、別のクライアント1bを利用して合成音声聞くことができる。

《実施例8》図25は、実施例8のクライアントとサーバとの構成を示す図である。この実施例8は、LANに辞書検索サーバ5と音声合成サーバ2とが接続され、辞書検索サーバ5で検索された単語の発音を音声合成サーバ2で音声合成して、辞書検索クライアント4に単語の辞書情報と共に単語の発音を音声出力する例を示している。

【0124】図25において、辞書検索クライアント4は、検索すべき単語を入力する単語入力部41と、入力された単語を辞書検索サーバ5に送信する単語送信部42と、音声波形データを受信する波形受信部13と、検索された単語情報を受信する辞書情報受信部44と、辞書情報受信部44で受信された辞書情報を表示する辞書情報表示部45と、波形受信部43で受信された符号化された波形データを復号する波形復号化部14と、復号されたデジタルの波形データをアナログの音声波形に変換するDA変換部151と、アナログ音声波形の高周波成分をカットするローパスフィルタ152と、ローパスフィルタ152を通過したアナログ音声波形を音声として出力するスピーカ153とで構成されている。

【0125】辞書検索サーバ5は、辞書検索クライアント4から送信される検索単語を受信する単語受信部51と、単語辞書53から検索すべき単語の意味及び単語の発音を示す表音記号等からなる辞書情報を検索する辞書検索部52と、辞書検索部52で検索された表音記号を音声合成サーバ2に送信する表音記号送信部55と、辞書検索部52で検索された辞書情報を辞書情報クライアント4に送信する辞書情報送信部54と、音声合成サーバ2から波形データを受信すると共に、受信した波形データを辞書検索クライアント4に送信する波形受信/送信部56とで構成されている。

【0126】音声合成サーバ2の構成は、基本的には前述した実施例と同じであり、この実施例では、データ受信部として辞書検索サーバ5から送信される表音記号を受信する表音記号受信部220が設けられている。

【0127】次に、図25のシステムの動作を説明する。ユーザにより単語入力部41から検索すべき単語が40 入力されると、入力された単語が単語送信部42から辞書検索サーバ5へ送信される。

【0128】辞書検索サーバ5では、単語検索部52が、単語受信部51で受信された単語の情報を単語辞書53から検索し、検索した情報の中で単語の意味などを示す辞書情報を辞書情報送信部54へ出力し、単語の発音を示す表音記号列を表音記号送信部55へ出力する。そして、表音記号送信部55が、単語の表音記号列を音声合成サーバ2へ送信する。

【0129】音声合成サーバ2では、表音記号受信部2

20が、辞書検索サーバ5から送信される表音記号列を受信すると、音響パラメータ生成部22が、その表音記号列に基づいて音響パラメータ時系列を生成し、波形生成部24が、生成された音響パラメータ時系列に基づいて音声波形を生成する。そして、波形符号化部25がその音声波形を符号化し、波形送信部26がその符号化された波形データを辞書検索サーバ5に送信する。なお、音声波形の圧縮は、ADPCM等の公知の波形符号化方式を用いる。

10 【0130】辞書検索サーバ5では、波形受信/送信部56が、音声合成サーバ2から送信される波形データを受信すると、辞書情報送信部54に波形データの送信と同期して辞書情報を送信するように指示する。

【0131】辞書検索クライアント4では、波形受信部13が、辞書検索サーバ5から送信される波形データを受信すると、波形復号化部14が、その符号化された波形データを復号し、さらにDA変換部151がデジタル音声波形をアナログ音声波形に変換する。DA変換部151で変換されたアナログ音声波形は、ローパスフィルタ152を通りスピーカ153から単語の発音として20 して音声出力される。同時に、辞書情報受信部44で受信された単語の辞書情報が、辞書情報表示部45に表示される。

【0132】なお、音声合成サーバ2で生成された音声波形を辞書検索サーバ5を経由して辞書検索クライアント4に送信しているのは、辞書検索サーバ5で検索される辞書情報と、音声合成サーバ2で合成された合成音声と同時にユーザに呈示する為である。また、図25には示していないが、辞書検索サーバ5に辞書情報バッファ及び波形データバッファを設け、一度検索した単語の辞書情報及び波形データを一定期間記憶しておくようにすれば、次に同じ単語の検索要求があった場合に、再度、単語の検索、音声合成処理を行う必要がなくなるので、検索結果をユーザに呈示するまでの時間を短縮できる。

【0133】この実施例では、検索要求のあった単語の辞書情報と共にその単語の発音を音声で出力するようにしたので、ユーザは、単語の意味と同時に発音を正確、かつ簡単に知ることができる。また、単語辞書53に格納されているそれぞれの単語の表音記号に基づいて、音声合成サーバ2が音声を合成するようにしたので、辞書検索サーバ5の単語辞書53に単語毎に音声データを記憶しておく必要がなくなり、音声合成の為に必要なデータを記憶しておく記憶装置の容量を少なくできる。

【0134】なお、音声出力するのは、1つの単語に限らず複数の単語からなる句あるいは文の発音を音声出力するようにしても良い。

《実施例9》次に、図26は、本発明の実施例9を示す図である。この実施例9は、スケジュール管理サーバ7で管理されるスケジュール情報を音声でユーザに知らせるものである。以下、前述した実施例と共通する部分に50

は同じ符号を付けてそれらの説明を省略する。

【0135】スケジュール管理サーバ7では、制御部71が、一定時間間隔でスケジュール検索部72を制御して、スケジュールデータベース73に登録されているスケジュール情報を検索し、現在時刻と一致するスケジュール情報が登録されているか否かを調べる。現在時刻と一致するスケジュール情報が登録されている場合には、そのスケジュール情報をスケジュール送信部74から音声合成サーバ2に送信させる。

【0136】音声合成サーバ2では、テキスト受信部21が、スケジュール管理サーバ7から送信されてくるスケジュール情報（テキストデータ）を受信すると、先ず、表音記号生成部22が、受信したテキストデータに基づいて表音記号列を生成し、次に音響パラメータ生成部23が、表音記号列に基づいて音響パラメータ時系列を生成する。さらに、波形生成部24が、その音響パラメータ時系列に基づいて音声波形を生成し、波形符号化部25がその音声波形を符号化して波形送信部26に出力する。そして、波形送信部26が符号化された波形データをスケジュール管理サーバ7に送信する。

【0137】スケジュール管理サーバ7では、波形受信／送信部75が音声合成サーバ2から波形データを受信すると、スケジュール送信部74と同期を取ってスケジュール情報と共に波形データをスケジュール管理クライアント6に送信する。

【0138】スケジュール管理クライアント6では、スケジュール受信部61がスケジュール管理サーバ7からのテキストデータを受信すると、スケジュール表示部62にそのテキストデータを表示する。同時に、波形受信部13で受信されたスケジュール内容を示す波形データが波形復号化部14で復号され、D/A変換部151でD/A変換されてスピーカ153から音声として出力される。

【0139】なお、スケジュール管理などのように、情報の表示とその情報の音声出力とを必ずしも同期させる必要がない場合には、スケジュール管理サーバ7を経由せず直接スケジュール管理クライアント6に音声波形データを送信するようにしてもよい。

【0140】この実施例によれば、スケジュール情報を文字情報としてばかりでなく、音声情報としても出力できるので、ユーザに対するスケジュールの通知をより確実に行うことができる。

【0141】

【発明の効果】本発明によれば、音声合成サーバから符号化した音声波形を送出し、クライアント側でその符号化された音声波形を復号することにより、LANに送出されるデータを減らし、LANのトラヒックを減少させることができる。

【0142】また、クライアントが表音記号生成機能、あるいは音響パラメータ生成機能を有する場合に、それ

ぞれのクライアントが自己の持つ機能に応じたデータを送信し、音声合成サーバ側で受信したデータの種類に応じた処理を行うことで、音声合成サーバの負荷を減らすことができる。さらに、LANの負荷状況、音声合成サーバの負荷状況等に応じてクライアントから音声合成サーバへ送信するデータの種類の、あるいは音声合成サーバからクライアントへ送信するデータの種類を変化させることで、クライアントと音声合成サーバの負荷分担を適正にして、システム全体の効率を良くすることができる。

【0143】さらに、音声合成サーバで生成された音声波形を、D/A変換機能を有する他のクライアント又は指定した電話機に送信させることで、D/A変換機能を有しないクライアントのユーザが合成音声を聞くことができる。

【0144】また、検索した単語の辞書情報と共にその単語の発音を音声で出力させることで、単語の発音を正確かつ簡単に知ることができる。さらに、スケジュール情報を音声で出力することで、ユーザに対するスケジュール内容の通知を確実に行うことができる。

【図面の簡単な説明】

【図1】本発明の原理図1である。

【図2】本発明の原理図2である。

【図3】本発明の原理図3である。

【図4】本発明の原理図4である。

【図5】本発明の原理図5である。

【図6】本発明の原理図6である。

【図7】本発明の実施例1を示す図である。

【図8】略字及び数字とその読みの対応テーブルを示す図である。

【図9】ADMを示すブロック図である。

【図10】ADPCMを示すブロック図である。

【図11】APCを示すブロック図である。

【図12】SBCを示すブロック図である。

【図13】ATCを示すブロック図である。

【図14】PARCORを示すブロック図である。

【図15】MPCを示すブロック図である。

【図16】VQを示すブロック図である。

【図17】本発明の実施例2を示す図である。

【図18】本発明の実施例3を示す図である。

【図19】本発明の実施例4を示す図である。

【図20】実施例1～4のクライアントを備えたシステム図である。

【図21】本発明の実施例5を示す図である。

【図22】実施例5の具体的構成の一例を示す図である。

【図23】本発明の実施例6を示す図である。

【図24】本発明の実施例7を示す図である。

【図25】本発明の実施例8を示す図である。

【図26】本発明の実施例9を示す図である。

27

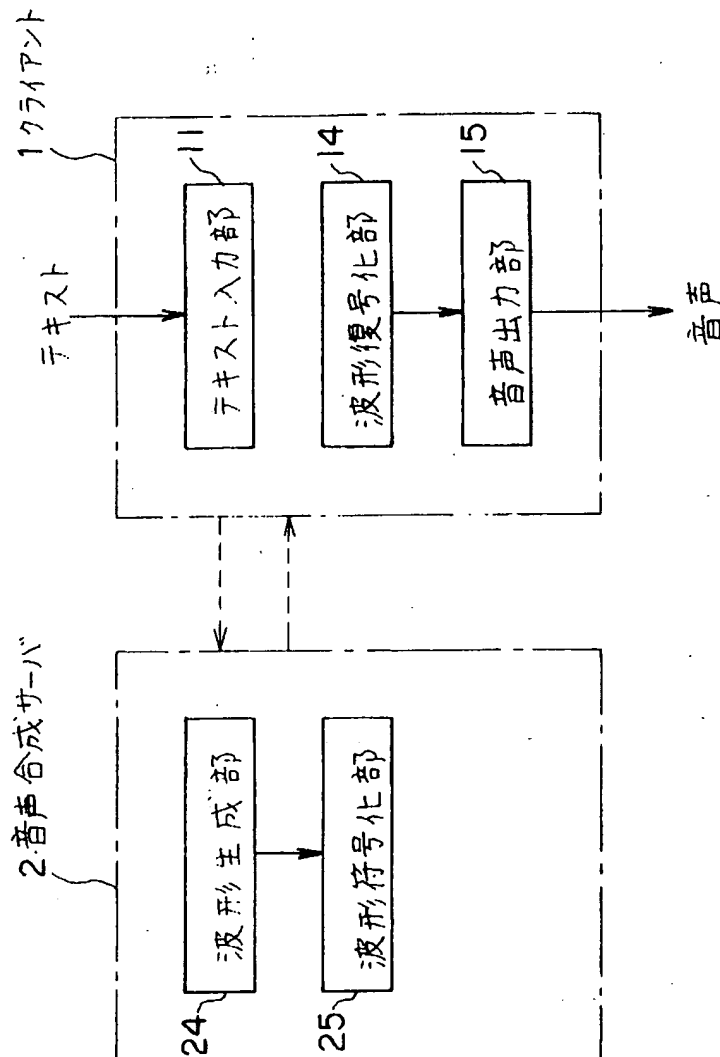
【図27】従来例を示す図である。

【符号の説明】

- 1、4、6 クライアント
 2 a、2 b 音響パラメータサーバ
 2 c 波形生成サーバ
 2 音声合成サーバ
 5 辞書検索サーバ
 7 スケジュール管理サーバ
 11 テキスト入力部

【図1】

本発明の原理図1



28

- 14 波形復号化部
 15 音声出力部
 16、22 表音記号生成部
 17、23 音響パラメータ生成部
 18、24 波形生成部
 25 波形符号化部
 110、260 データ送信部
 120、210 データ受信部
 130、270 データ制御部

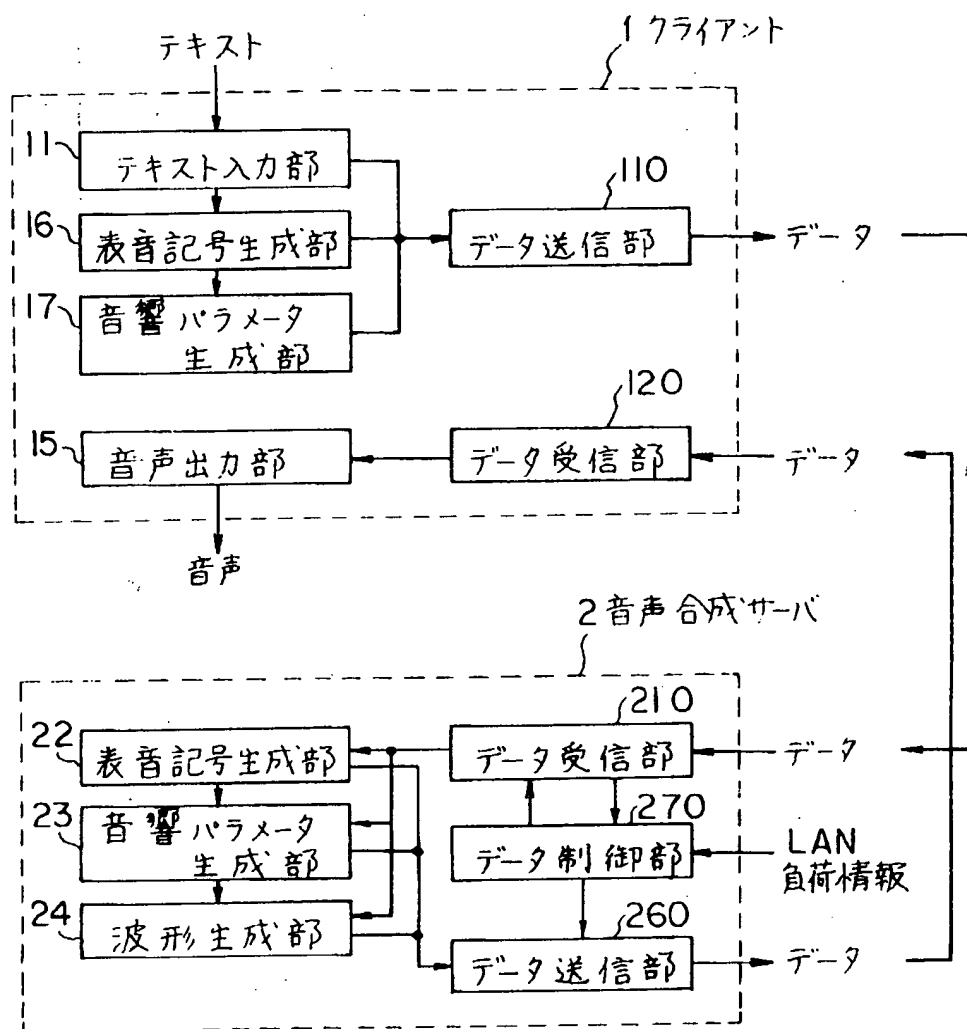
【図8】

略字・数字とその読みへの対応テーブル
 を示す図

テキスト中の略字・ 数字等	読み
IMF	あいえいふ
1991	せんきゅうひゃくきゅうじゅういち
16:25	じゅうろくにじにじゅうごふん

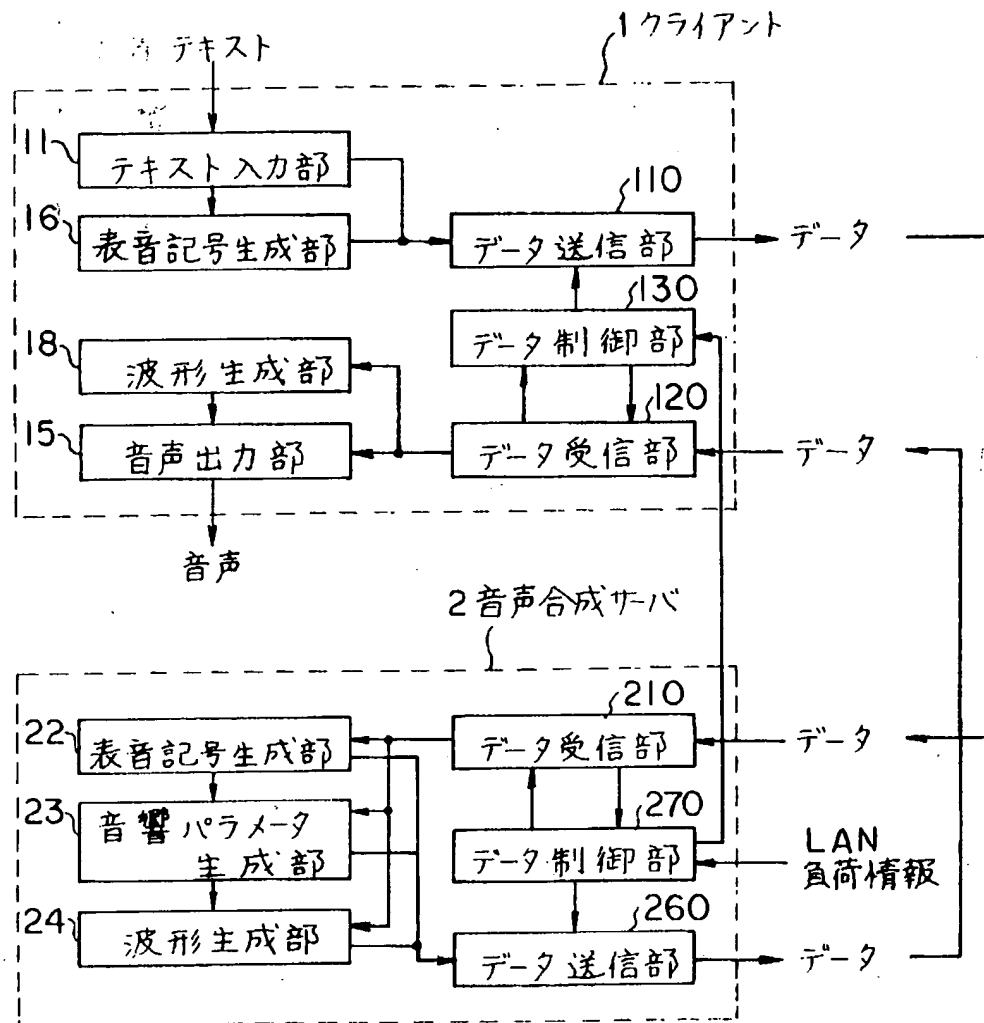
【図2】

本発明の原理図2



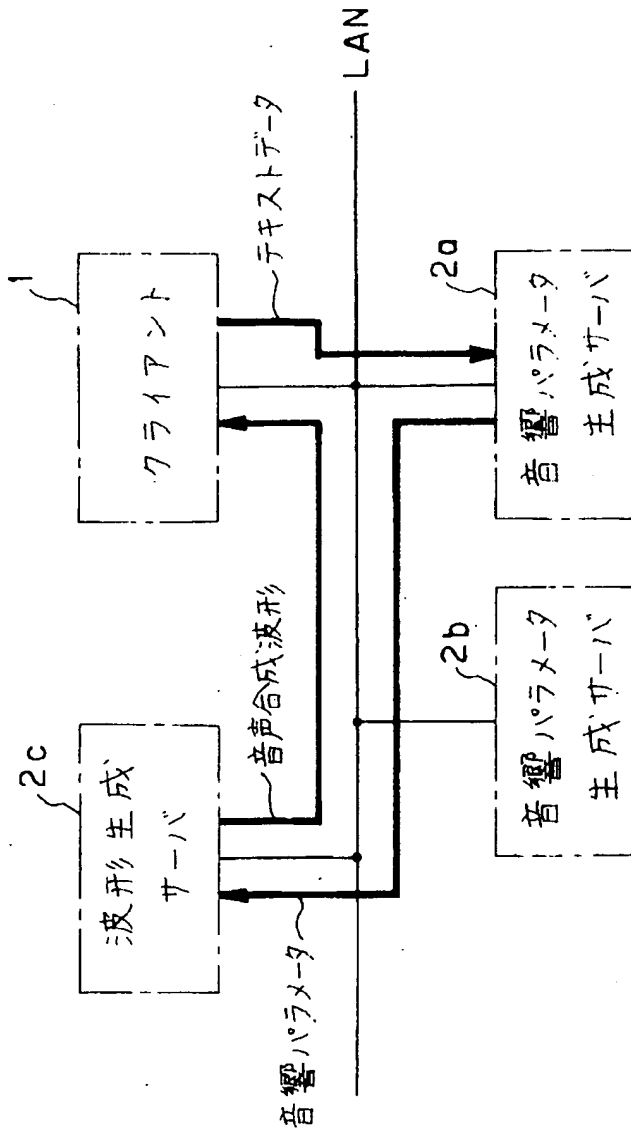
【図3】

本発明の原理図 3



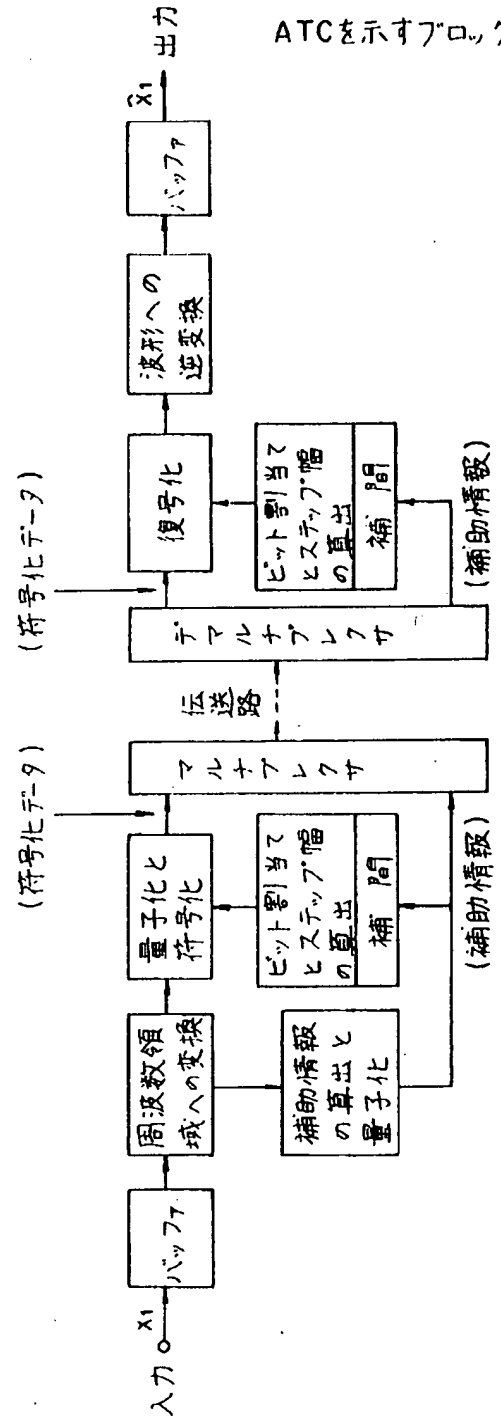
【図4】

本発明の原理図4



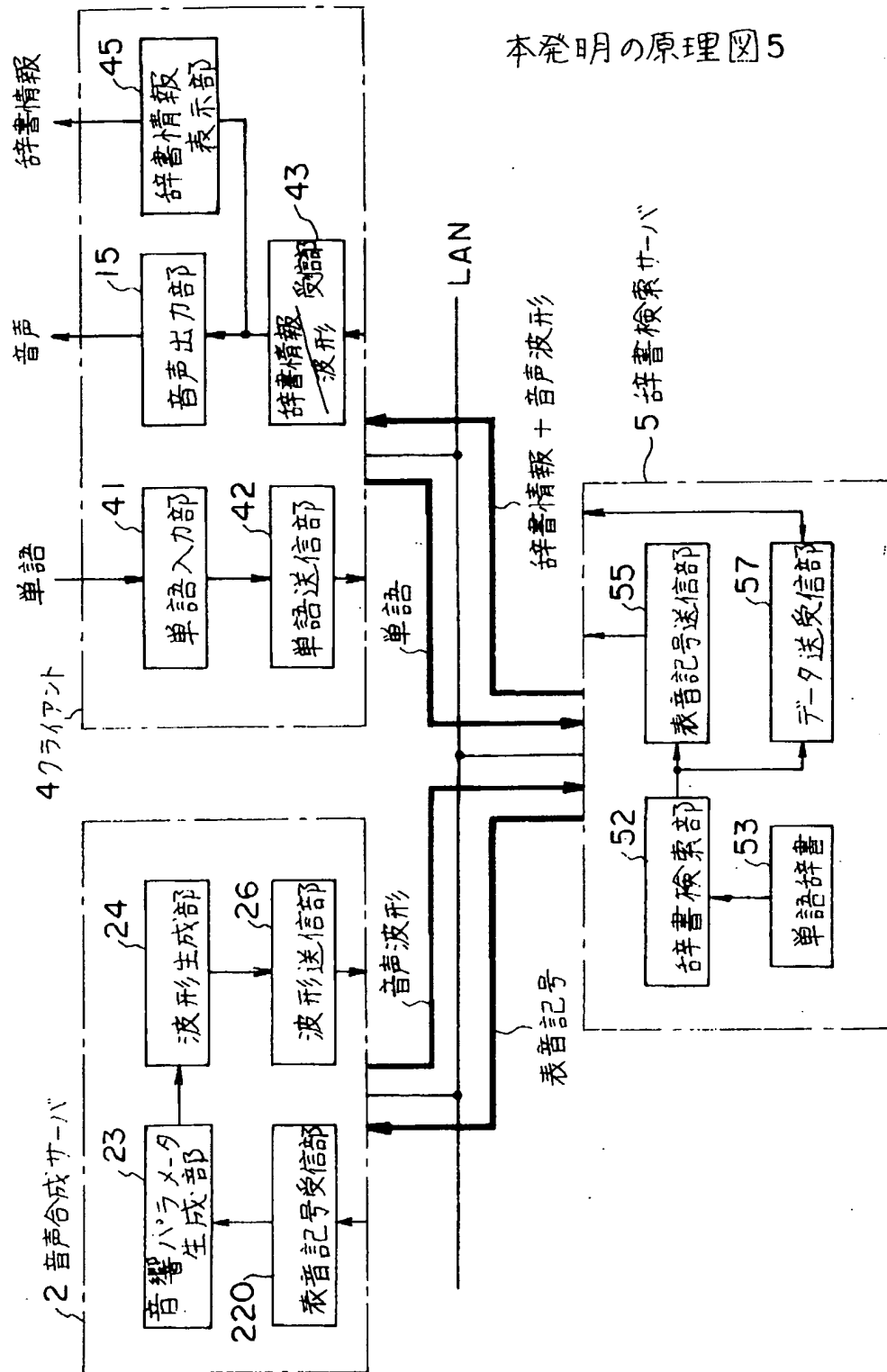
【図13】

ATCを示すブロック図

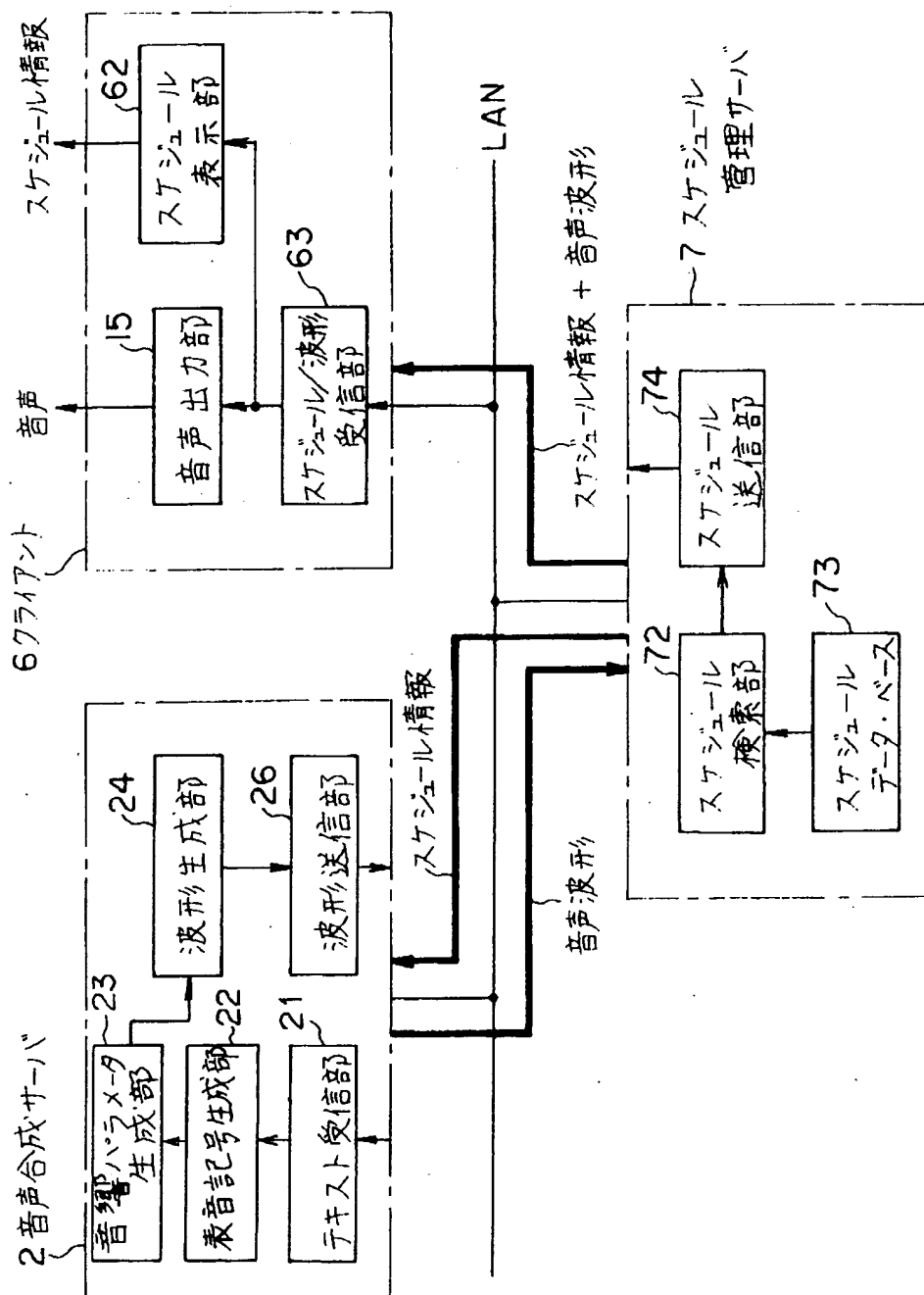


【図5】

本発明の原理図5

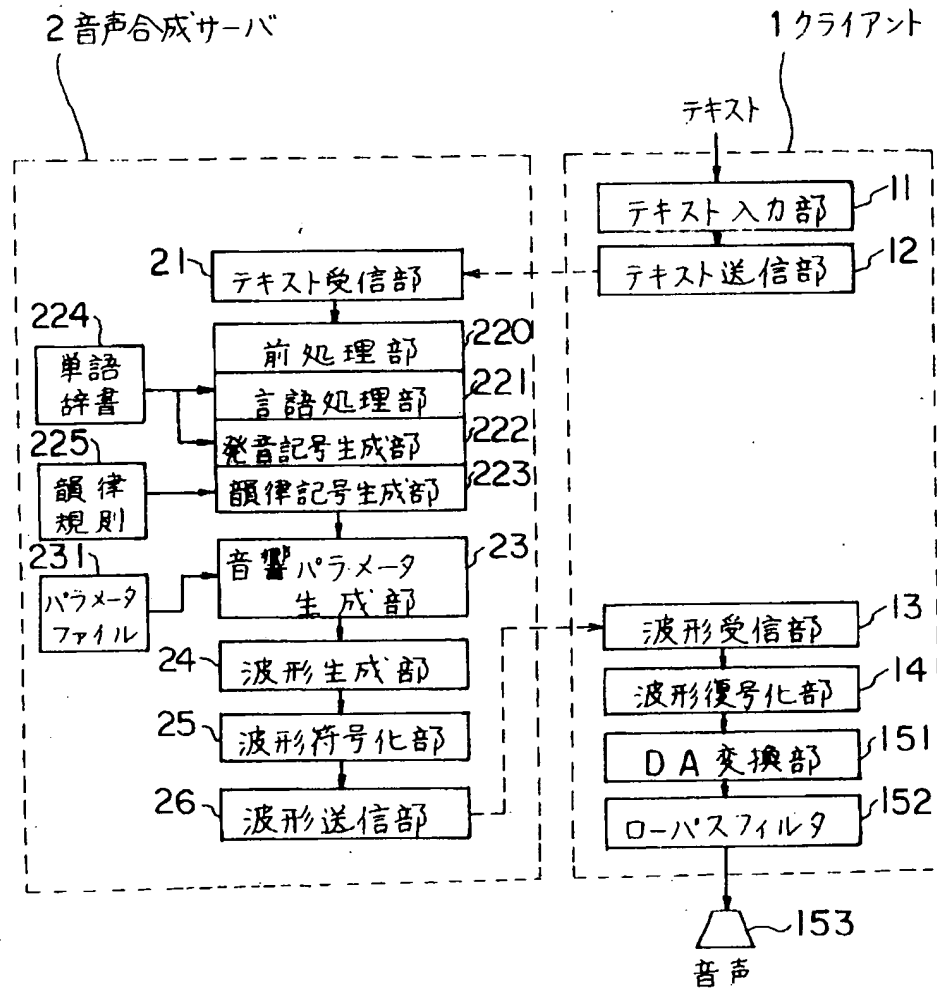


本発明の原理図 6



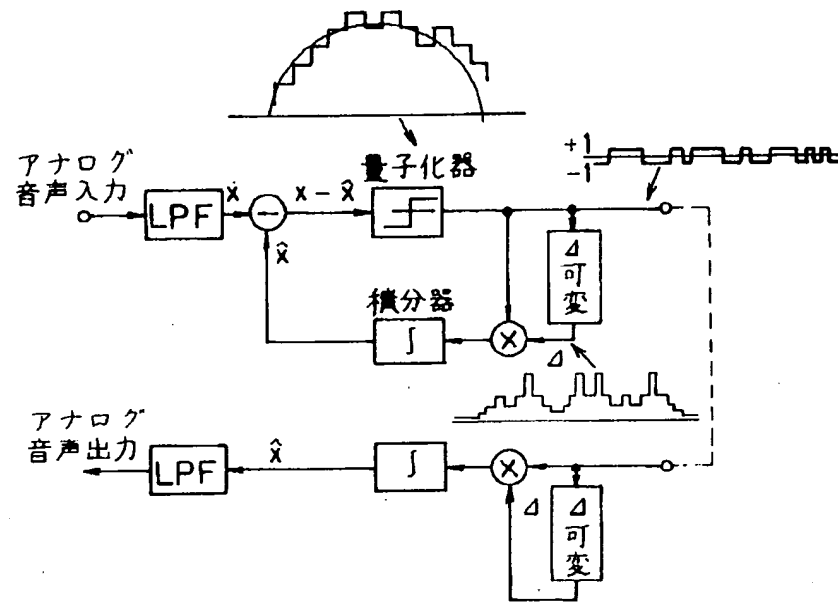
【図7】

本発明の実施例1を示す図



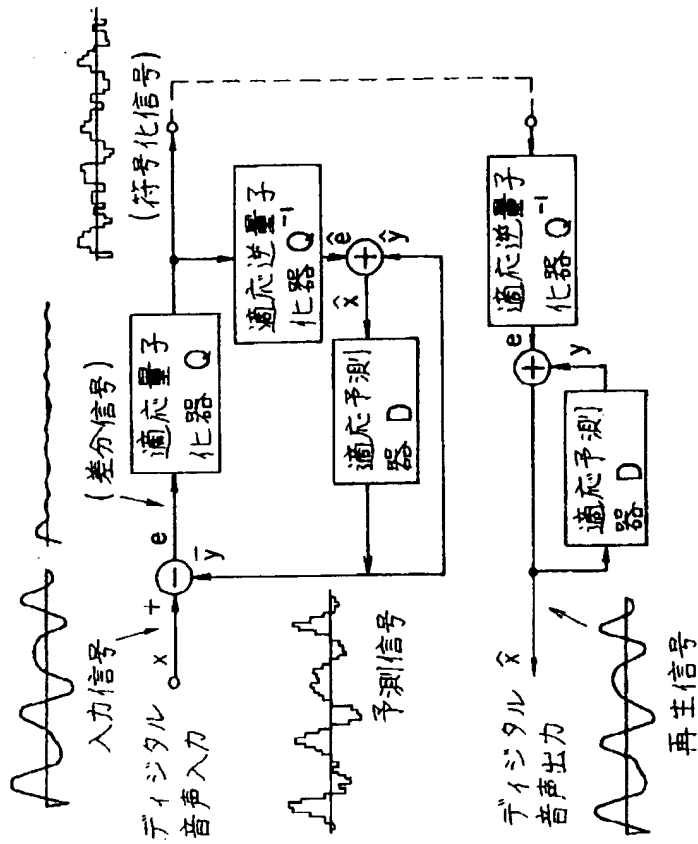
【図9】

ADMを示すブロック図



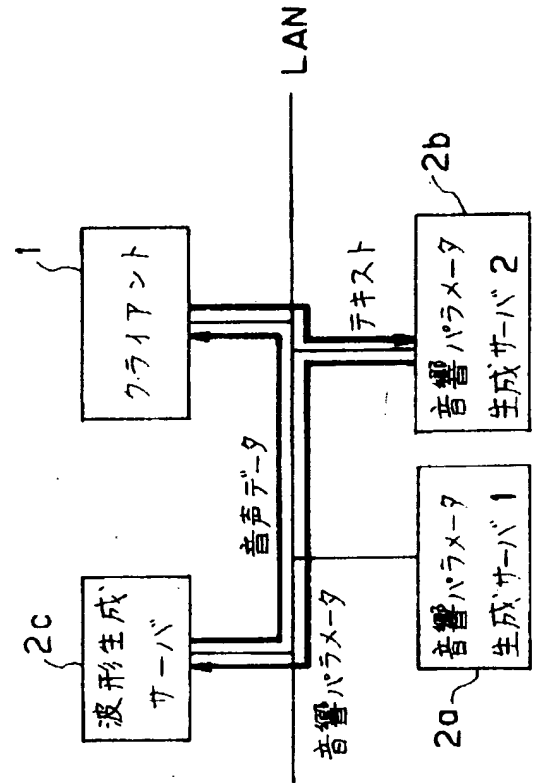
【図10】

ADPCMを示すブロック図



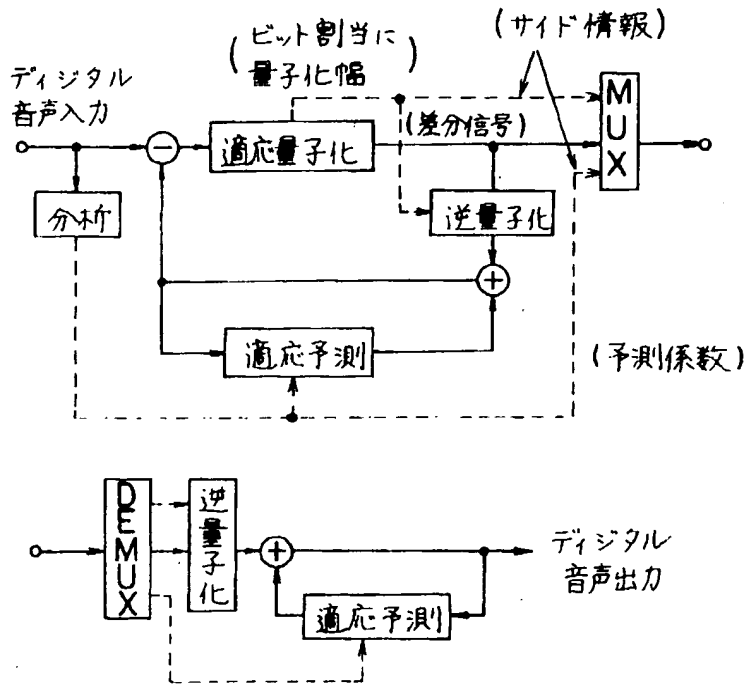
【図21】

本発明の実施例5を示す図



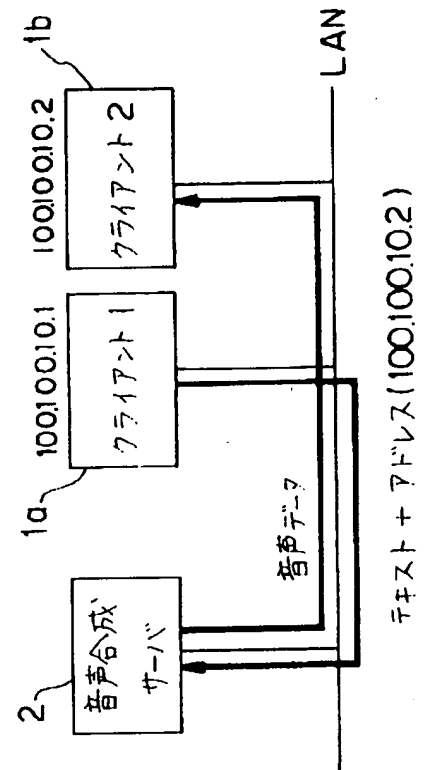
【図11】

APCを示すブロック図



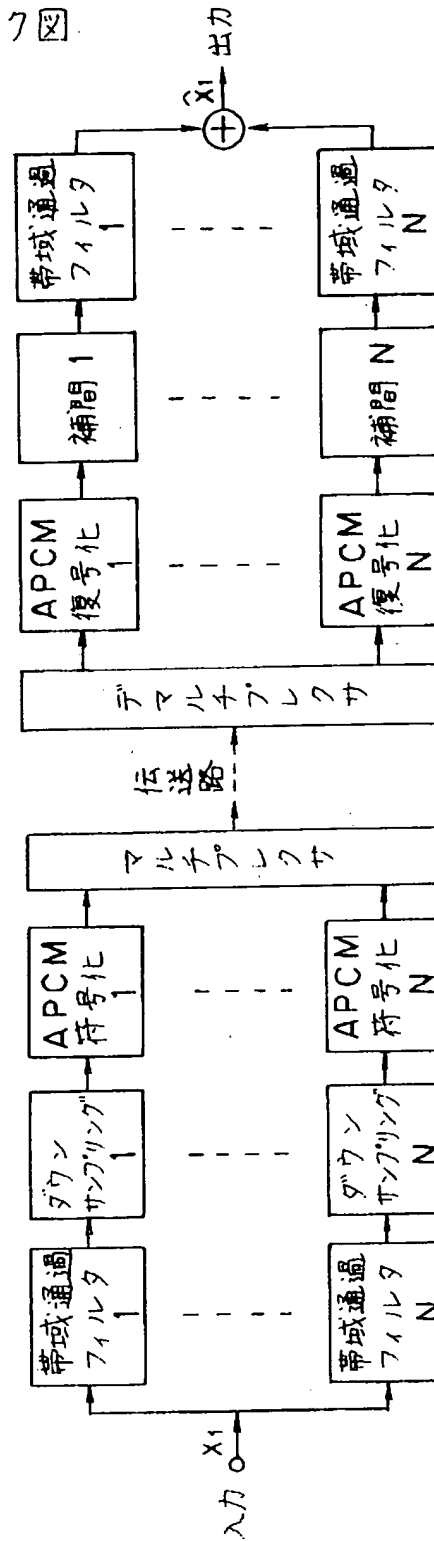
【図24】

本発明の実施例7を示す図



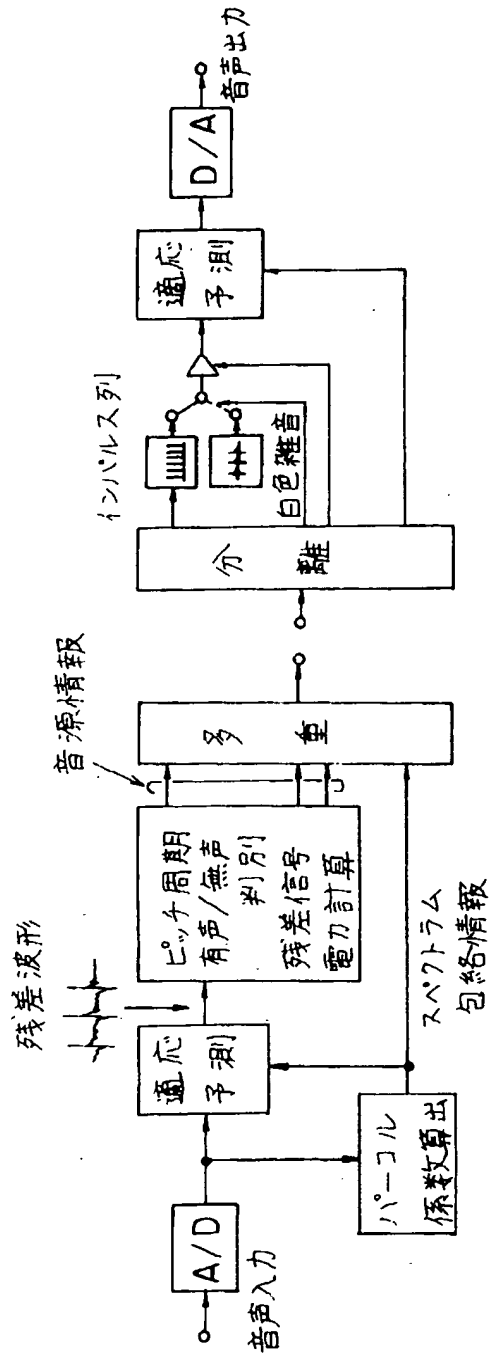
【図12】

SBCを示すブロック図



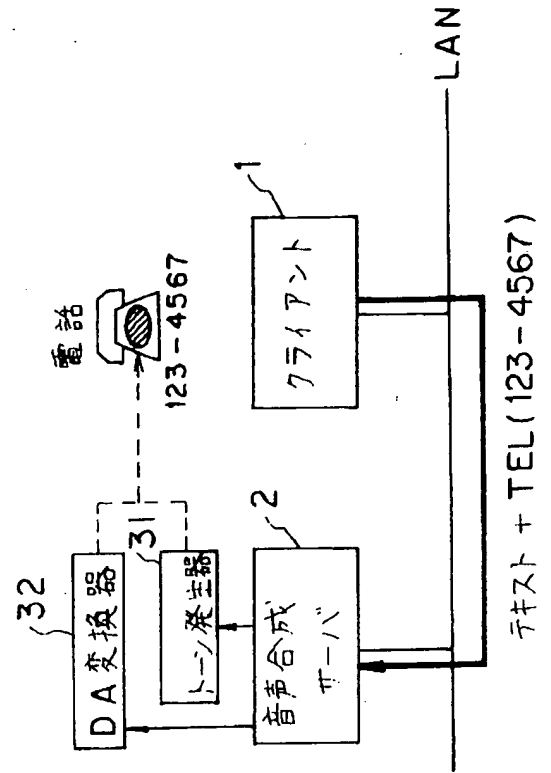
【図14】

PARCORを示すブロック図



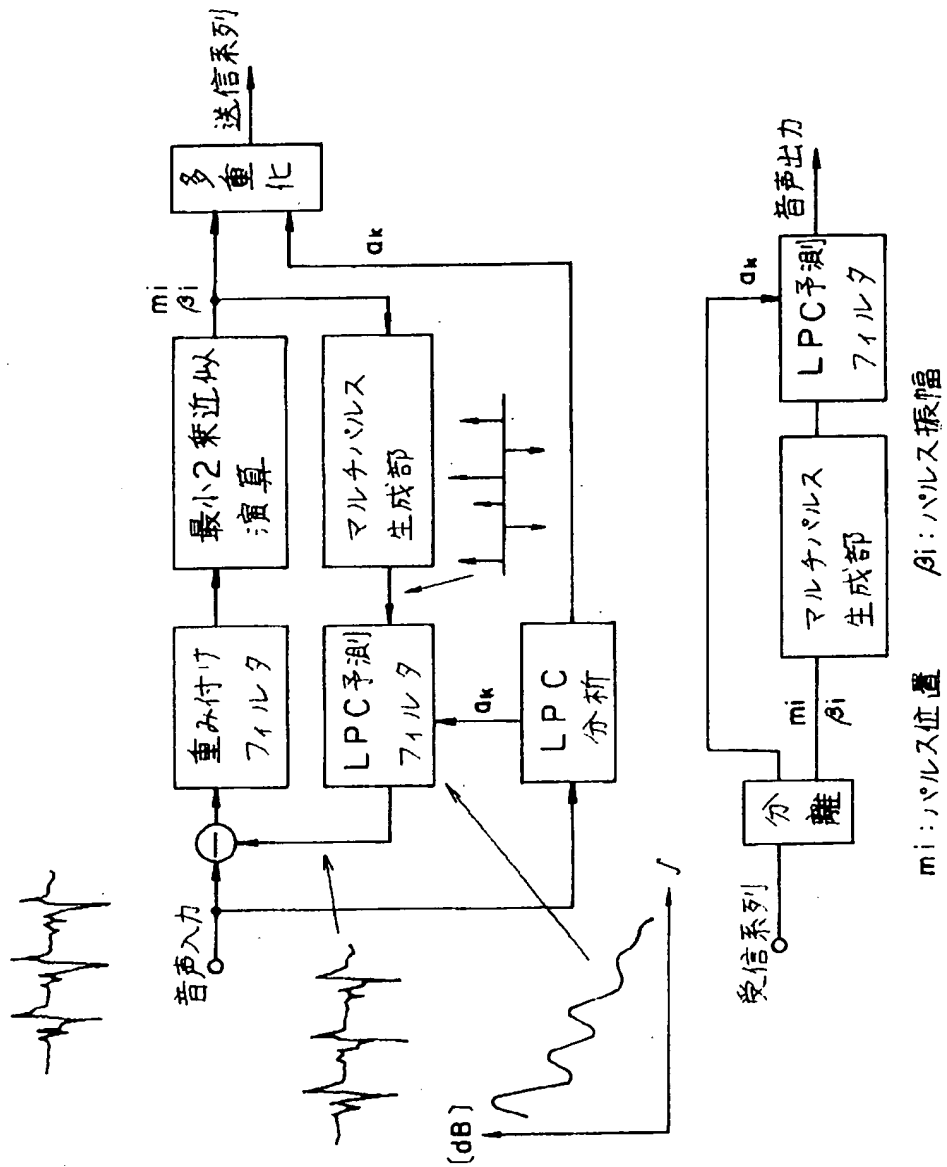
【図23】

本発明の実施例6を示す図



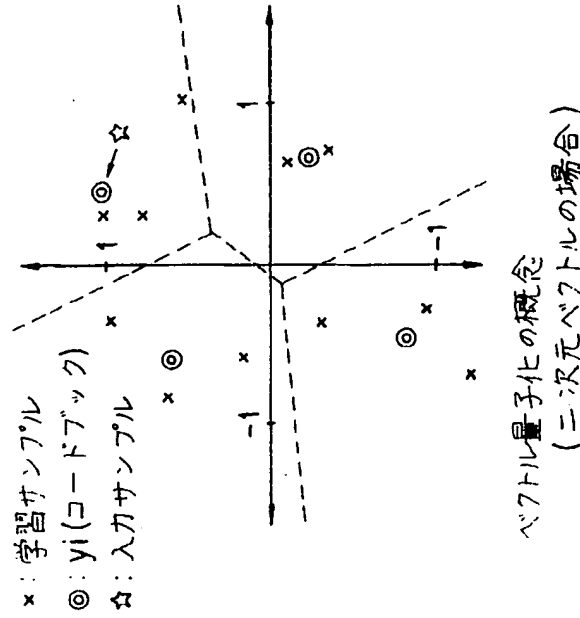
【図15】

MPCを示すブロック図

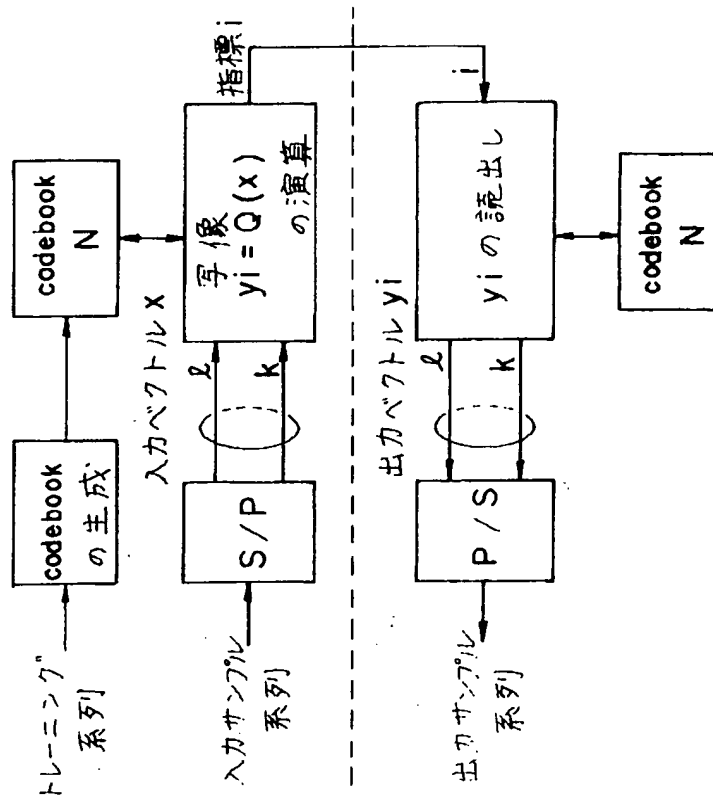


【図16】

VQを示すブロック図

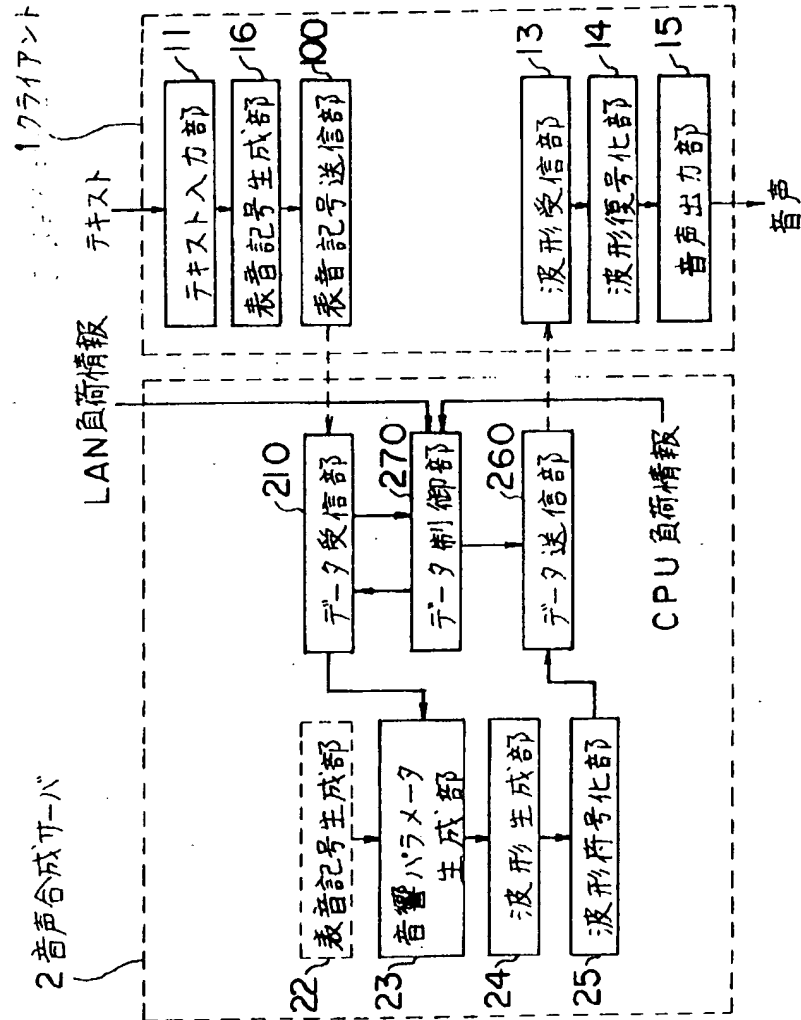


- 9.6 kbit/s 以下で用いられる。
- ベクトル辞書が必要 (メモリ最大)
- 良好な品質 (PCM 5~6 ビット相当)



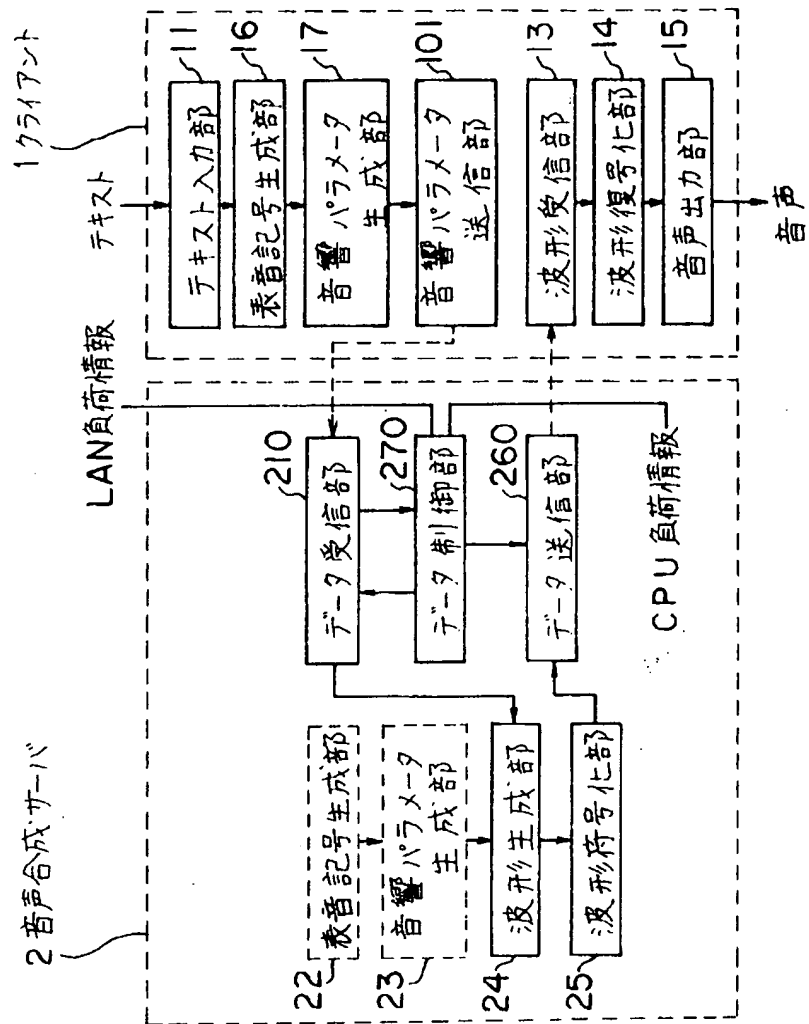
【図17】

本発明の実施例2を示す図



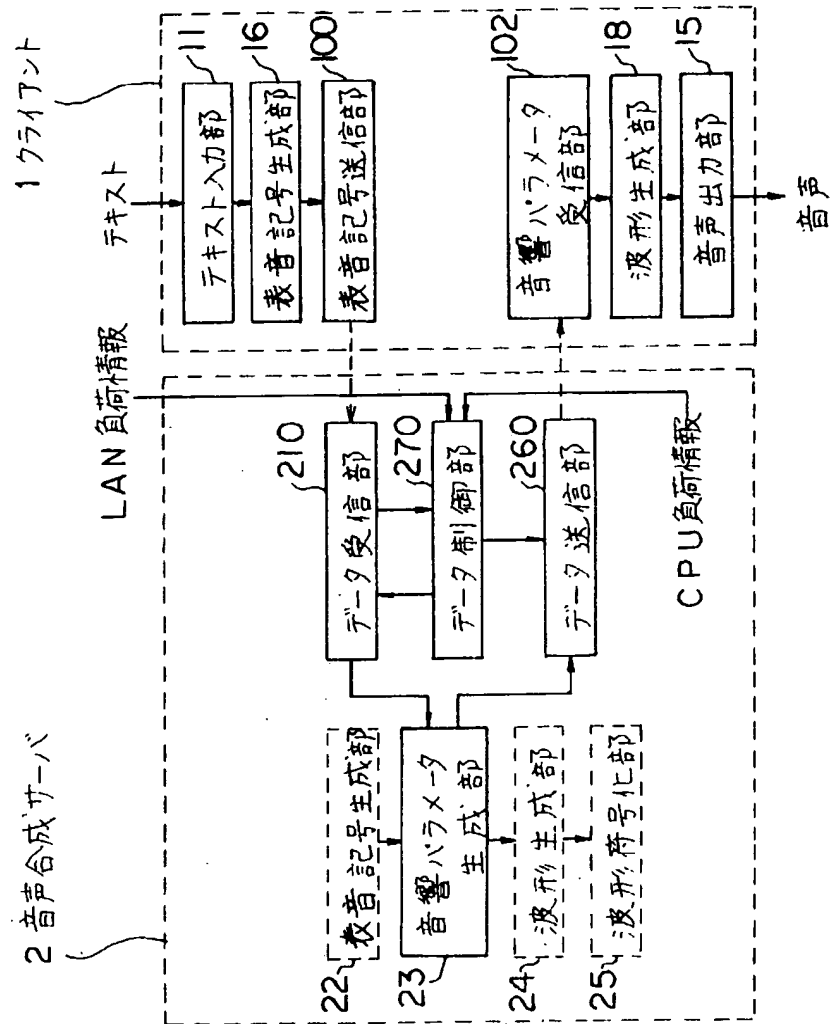
【図18】

本発明の実施例3を示す図



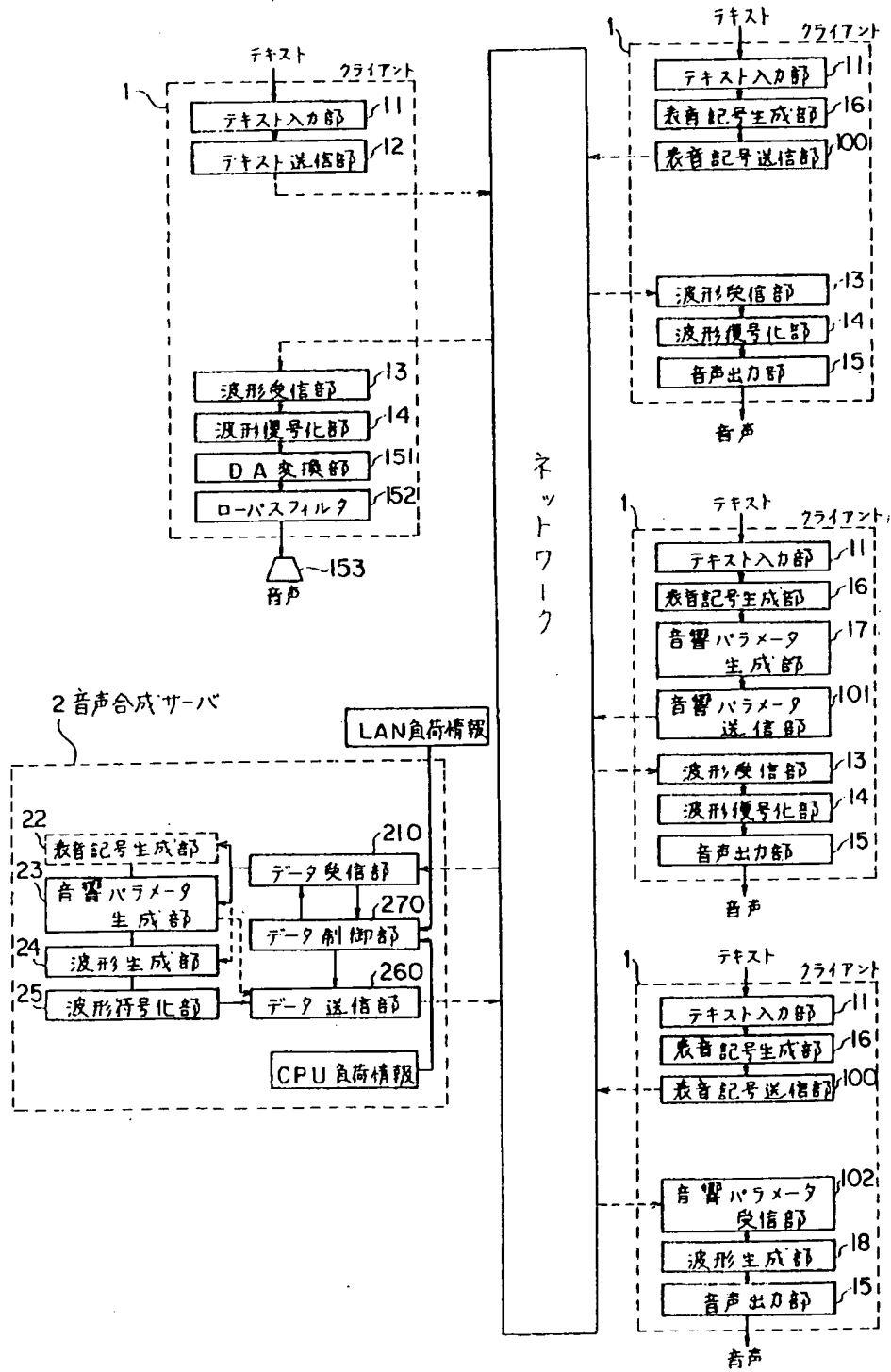
【図19】

本発明の実施例4を示す図



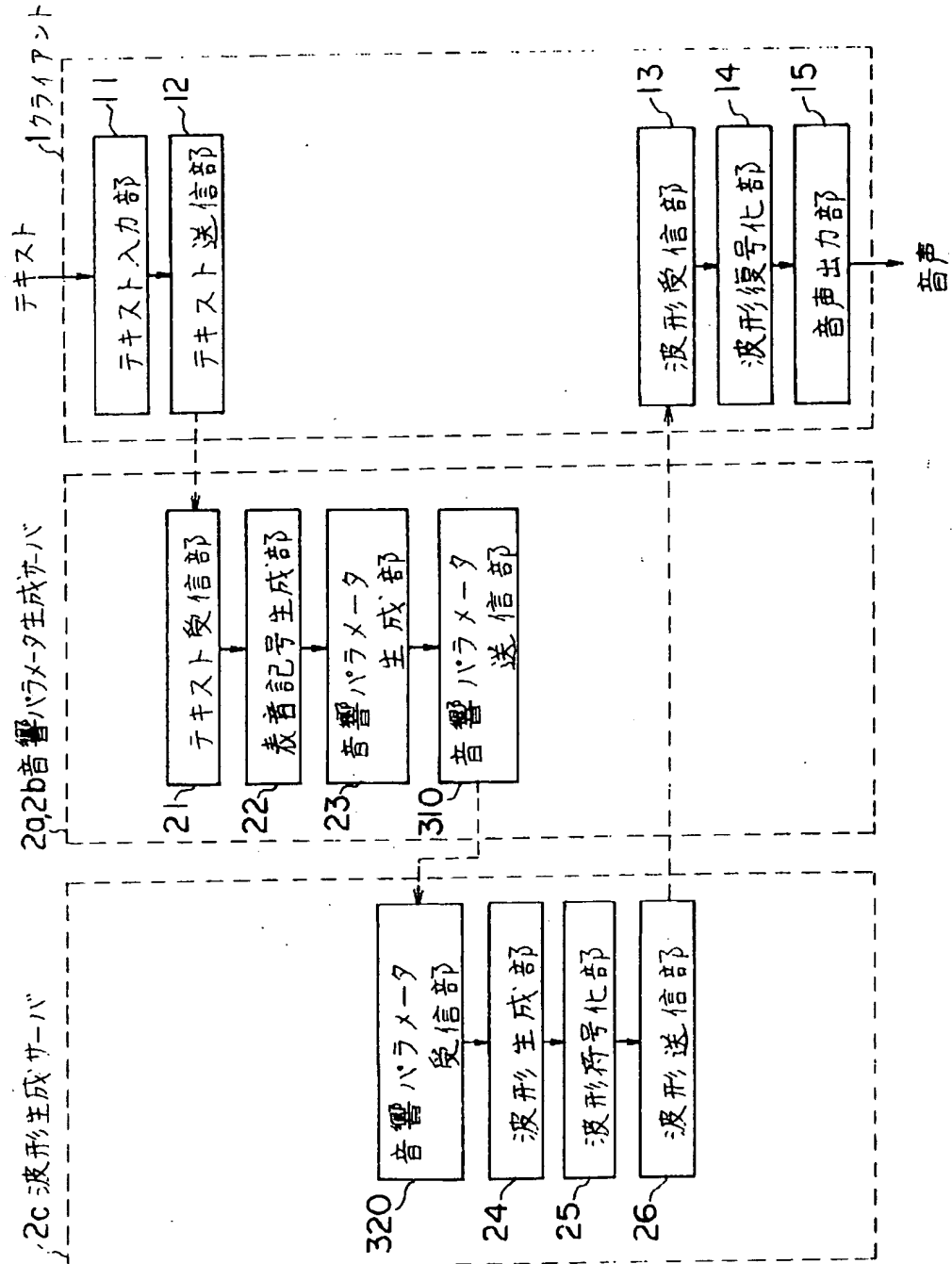
【図20】

実施例1,2,3,4のクライアントと備えたシステム構成図



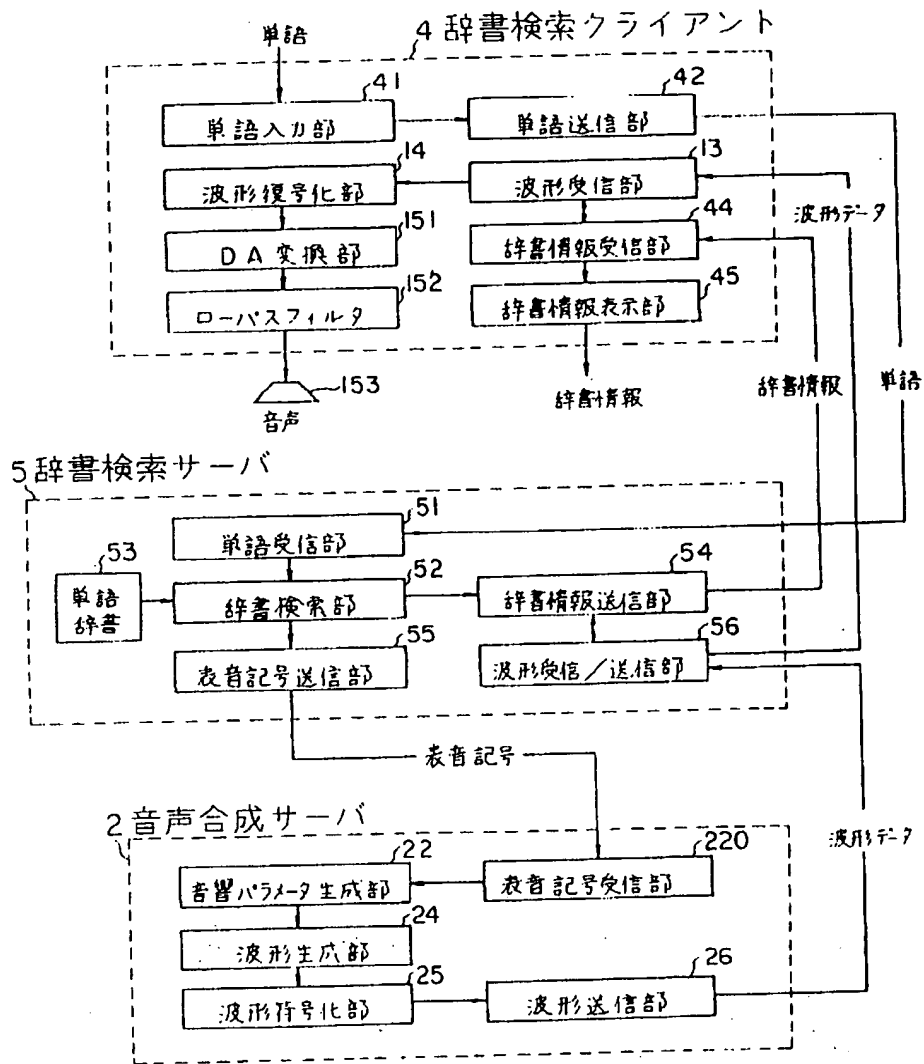
【図22】

実施例5の具体的構成の一例を示す図



【図25】

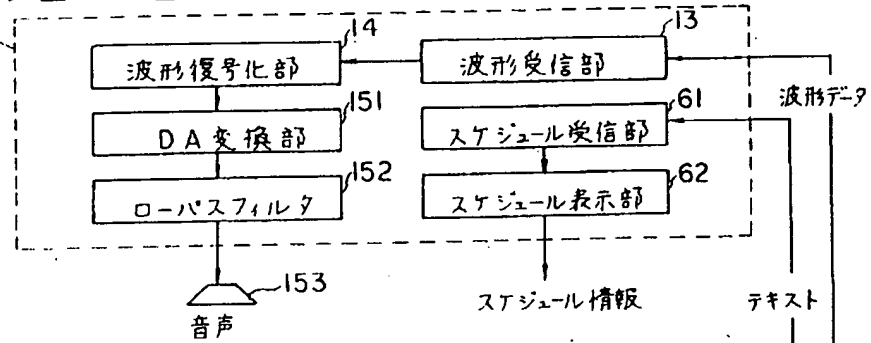
本発明の実施例8を示す図



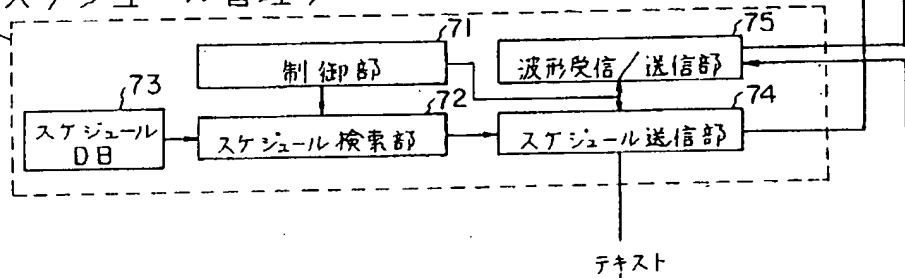
【図26】

本発明の実施例9を示す図

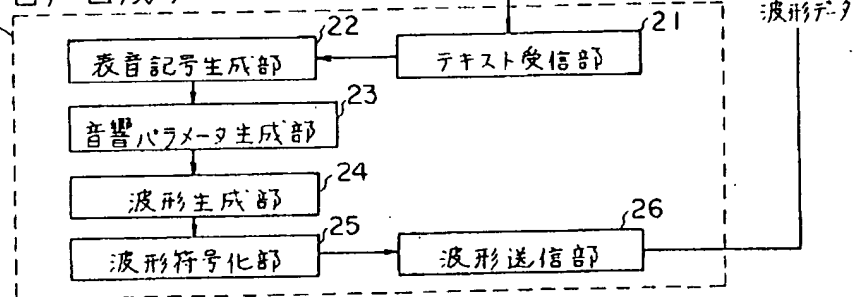
6 スケジュール管理クライアント



7 スケジュール管理サーバ



2 音声合成サーバ



【図27】

従来例を示す図

